University of Dundee

MASTER OF FORENSIC ANTHROPOLOGY

Ancestry determination using geometric morphometrics

Gillick, Hana

Award date:
2012

MASTER OF SCIENCE

Ancestry determination using geometric morphometrics

Hana Gillick

2012

University of Dundee

ANCESTRY DETERMINATION USING GEOMETRIC MORPHOMETRICS

HANA GILLICK

MRes in Forensic Anthropology

University of Dundee

August 2012

TABLE OF CONTENTS

FIGURES

TABLES

ACKNOWLEDGEMENTS

My most sincere gratitude goes towards my supervisors, Professor Caroline Wilkinson and Dr. Andrea Cardini. I am grateful for their help, support, and encouragement throughout this research, both on an academic level and on a personal level. A really special thank you to Dr. Andrea Cardini for his patience and for providing me with so much of his time and advice on how to develop this project and make it better.

Thank you to Dr. Gordon Findlater for allowing me access to the collection of skulls in the Anatomy Museum at the University of Edinburgh.

I would also like to thank my fellow graduate students and the staff members from CAHID for their encouragement and suggestions and for being there and being themselves. I am truly grateful to have met such wonderful, interesting, and caring people.

Above all, I would like to thank my parents for their support, for so strongly believing in me, and pushing me. And thank you for having such a genuine interest in my research.

DECLARATION

The candidate is the author of this thesis. Unless otherwise stated, all references cited have been consulted by the candidate. The work, of which the thesis is a record, has been completed by the candidate. No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university, or institute of learning.

Hana Gillick

ABSTRACT

Ancestry identification is an important aspect in forensic anthropological investigations concerning the biological and social identity of unidentified human remains. Traditionally, ancestry is usually assessed by the visual examination of the shape and form of the features of the skull and/or measured by linear distances between certain cranial landmarks. In this research, three-dimensional Procrustes based geometric morphometric methods are employed to investigate whether these techniques can classify ancestry in a forensic context using the cranium. Cranial landmark and semilandmark configurations, focused specifically on features routinely used to make ancestry assessments using traditional morphological and morphometric techniques, are used to classify ancestry using discriminant function analyses with leave-one-out cross validation. Landmarks and semilandmarks were captured on 3D cranial images of 31 sub-Saharan African and 31 European individuals and variables of size, shape, and form were used to assign crania into respective groups based on different configurations of points. Results demonstrate that variables of shape and/or form (i.e. size and shape) are more accurate at classifying ancestry than size alone, and that an overall aspect of the cranium is more accurate at classifying individuals into their respective ancestry groups than semilandmarks recorded along specific features, such as the border of the orbits, the edge of the lower nasal aperture, and along the vault on the midplane.

1. INTRODUCTION

Ancestry/race assessment forms one of the four basic components of forensic anthropological casework in investigations concerning the biological and social identity of unidentified human remains. Traditional ancestry assessment techniques generally involve examining the shape and form of features of the skull through visual analysis and by the use of linear distance measurements. This research focuses on the application of modern methods of shape analysis to the process of ancestry classification using the cranium. Geometric morphometric techniques are used to explore cranial shape, size, and form differences in individuals with differing ancestral origins through the use of conventional cranial landmarks and semilandmark curve data. The cranial landmarks and semilandmarks used in this study are focused on those features that are routinely used to make ancestry assessments using traditional morphological and morphometric techniques. Subsets of shape and form coordinates and size variables, extracted from the landmark and semilandmark datasets using geometric morphometric and multivariate analyses, are examined to determine whether particular subsets of landmarks, representing certain cranial features, are more effective in correctly assigning individuals to their respective ancestry populations.

1.1. THE CONCEPT OF RACE

"There are few words employed in physical anthropology that stir more varied reactions than "race"" (Ubelaker, 1996:236). Physical anthropology has played a major role in the construction and dissemination of the concept of biological race (Harrison, 1995). The study of human variation and the classification of human biological differences formed the theoretical foundation of early physical anthropology (Caspari, 2003). Historically, physical anthropologists believed that human beings were

biologically divided from one another and could be classified naturally into a broad range of taxonomic racial groups. In the past, human races were seen as separate and immutable and were defined according to a distinct set of phenotypic and morphological descriptors wherein cultural, mental, and behavioural characteristics often played a part in the classificatory rubric (Caspari, 2003). Ethnocentric prejudices and hierarchical racial classifications infiltrated much of the early research in physical anthropology. The advancement of research into evolutionary biology and genetics, as well as the social sciences, led most anthropologists, biologists and many other academics to eventually reject the validity of the biological race concept. Even though the concept of biologic race has been largely discredited, the concept itself has been so pervasive that it continues to be an ever-present subject in society as well as in parts of academia (Kennedy, 1995). In general, however, it is widely accepted among academics that using the term race today to account for human biological variation is highly problematic and flawed (Kennedy, 1995; Caspari, 2003).

For many, use of the term "race" in scientific disciplines "harks back to previous centuries when overly simplistic and erroneous notions of human variation prevailed and racist attitudes worked their way into the scientific literature" (Ubelaker, 1996:236). Even though anthropology was, historically, a huge contributor to the proliferation of racial ideologies, the discipline has also played a major role in the de-construction and eventual re-construction of racial concepts as socially regarded concepts (Harrison, 1995). Although the term "race" was used to describe human phenotypic, morphologic and genetic variation in biological contexts in the past, most physical anthropologists make a deliberate attempt to avoid using the term today. Presently, terms such as "ancestry", "population affiliation" or "geographic origin" are more comfortably applied in physical anthropology in discussions concerning human group variation.

The concept of "race" is often an emotional and highly charged subject. Most anthropologists agree that it is a cultural construct used ambiguously in most societies to differentiate between people based on geography, political history, religion and socioeconomic status, and most often relates to externally visible characteristics such as skin colour, hair texture, and facial features (AAPA Statement on biological aspects of race, 1996; Bolnick, 2008). The less contentious term "ancestry" usually refers to the geographic region or regions from where an individual's biological ancestors originated, and is generally described, geographically or continentally, in terms of major worldwide populations such as African, European, Asian, and Australian (Klepinger, 2006; Bolnick, 2008). Ancestry does not necessarily denote race, although there appears to be a correlation between the terms both biologically and socially (Klepinger, 2006; Bolnick, 2008). The terms race and ancestry are sometimes used synonymously, especially within forensic anthropology and human identification, which can greatly confuse the matter (Klepinger, 2006:64).

It is clear that phenotypic differences exist since differences in skin colour, hair texture, body shape and facial features are apparent. But these visible differences do not clearly indicate that humans can be naturally divided into sets of discrete biological racial categories (Kennedy, 1995). Human variation is mainly seen as clinal and environmentally adapted, meaning that there are no obvious boundaries between people and that changes in human form and appearance are gradual as geographic distance between populations increases (Livingstone and Dobzhansky, 1962; Handley et al., 2007).

Although race as a biological or scientific concept has been mostly rejected, sociological concepts of race have persisted within cultures and societies. Socially

speaking, race is real and relevant (Cartmill, 1998). Individuals are organized into racial, ethnic, and ancestry categories on a regular basis in government census surveys, polls, the mass media, medicine, research studies, forensic databanks, and in legal documents and investigations (Kennedy, 1995). Racial concepts and constructs can be important aspects in establishing individuals' social and cultural identities in the societies in which they reside. Race can affect experiences and life opportunities and so membership in a racial group can play an important role in the way people view themselves and the others around them (Cartmill, 1998).

1.2. ANCESTRY/RACE AS AN IDENTIFIER IN FORENSIC ANTHROPOLOGY

Even though most physical anthropologists reject race as a sound scientific concept, "race identification continues to be one the central foci of forensic anthropological casework and research" (Sauer, 1992:107). When unidentified human remains are discovered, law enforcement agencies may contact practitioners in forensic anthropology and osteology for help in the recovery and analysis of the remains. When faced with decomposed or skeletonized human remains, the ultimate goal of practitioners working in forensic anthropology is to extract as much information as possible about the social and biological identity of the deceased. The primary objective of police and forensic practitioners involved in these scenarios is the individualization of the unidentified human remains.

The forensic anthropologist will attempt to establish a biological profile of unknown human remains. This consists of estimations of the age at death, sex, stature, and ancestry (or race) of the unidentified person or persons. Once police are provided with basic profile information they can begin to focus their search, for example, by attempting to match the forensic profile information with missing-person data on record.

Identification, specifically, of the ancestry/race of human remains may enable law enforcement agencies to focus their investigations on certain communities or groups within a more diverse society at large. Ancestry/race identification, therefore, even with all its attendant controversy and potential fallibility, is still considered a crucial part in many legal investigations and a practical means of identification within society. Indeed, perceived race may be as important of an aspect in individualizing skeletal remains as age at death and sex (Gill and Rhine, 1990; Kemkes, 2007).

There is a concordance between geographic ancestry and the socially conceived notion of race (Ousley, et al., 2009). Because race labels are in operation within most societies as a means of classifying/identifying individuals, forensic anthropologists tend to report information to law enforcement and the general public in a way that will be easily understood (Gill and Rhine, 1990). When forensic anthropologists are asked by law enforcement officials to determine the race of human skeletal remains they are usually attempting to identify ancestry/race based on "social" race or "bureaucratic" race concepts prevalent in the society in which the investigation is being carried out. The criteria used in the race identification process will rely on particular skeletal characteristics that are more frequently observed within certain populations due to different geographic microevolutionary origins (Albanese and Saunders, 2006). The ancestral or racial identification labels determined by the forensic examination will necessarily correspond with the broad categories and popular notions of race – such as African (or Black), European (or White) – as used by the law enforcement agencies or communities involved in the investigation (Albanese and Saunders, 2006). The justification for the use of racial identification in forensic evaluation is clear: any information that can be provided to police about how individuals may have looked or

where they may have lived during their lifetime will only help to improve the chances of identification.

The identification of ancestral origins from skeletal remains in a forensic context can be challenging. Ancestry classifications can be ambiguous and the methods used, at least traditionally, have often been subjective in nature (Albanese and Saunders 2006). Determining ancestry from distinguishing features and characteristics of skeletal remains is not as straightforward and not as standardized as some of the other aspects of the biological profile. Methods for establishing the age, sex, and stature of an individual from skeletal remains have been studied and scientifically tested to a greater extent. With regard to ancestry/race determination, Hefner states that there is a "lack of a methodological approach and, more importantly, the fact that there are no error rates associated with ancestry prediction…suggest that they have not been investigated with appropriate scientific and legal considerations in mind" (Hefner, 2009:985). An assessment of ancestry may, in some cases, be an important initial identifying characteristic since some of the methods used to determine age, sex, and stature require prior knowledge of ancestry in order to be applied with any degree of confidence (Byers, 2005).

The skull is considered to be the most useful part of the skeleton for the estimation of ancestry (Howells, 1973; White and Folkens, 2005; Klepinger, 2006; Randolph-Quinney et al., 2009). Specifically, the area of the facial region of the skull, is considered to be most indicative of variation between different geographic populations (Rhine, 1990:13). Some of this variation is easily detectable with the naked eye and one approach to ancestry determination is to visually inspect the skull for differences in shape and morphology (Hefner, 2009). This approach uses qualitative techniques and is

commonly called the non-metric approach. The quantification of these visible shape differences in the skull is important in a legal sense because it provides a less subjective method for studying craniofacial change. To be admissible in a courtroom, skeletal assessment techniques must be precise, objective, and quantifiable with known levels of confidence (Curran, 1990; Gill and Rhine, 1990).

A number of different studies have attempted to quantify the visible changes, based on ancestry, that are observed in the skull and craniofacial region. Although these metric methods are more objective than the non-metric approaches to ancestry determination, they are not without their limitations. These methods rely mostly on distance measurements taken with calipers between particular anatomical points of interest. These anatomical landmarks may sometimes be difficult to replicate and measurements of distances between various landmarks may not necessarily provide enough attention to important areas of ancestrally related morphological variance in the human skull (Rhine, 1990). The human skull is a complex and irregular structure not easily quantified by linear distances. Most cranial bones are highly curved with numerous concavities, projections and prominences that may be difficult or even impossible to quantify using traditional metric tools such as spreading and sliding calipers.

Geometric morphometric methods may provide a more efficient technique to assign quantities to biological shapes and forms than traditional morphometric and morphological approaches that have been traditionally applied (Bookstein, 1982). In general, geometric morphometrics is a field of study that utilizes a series of two- or three-dimensional co-ordinate landmarks on specimens to describe shape statistically and evaluate how it may vary between groups and specimens. These methods provide

quantitative information about size, shape, and shape change that can be analyzed statistically without losing vital information about the complex geometry of the specimen under study. Geometric morphometrics methods were primarily used in biological and medical contexts in the past. The methods have been adopted more recently into the field of physical anthropology. Although, to date, applications of geometric morphometrics in the field of forensic anthropology are few, practitioners in this field are now beginning to recognize the potential power of these statistical techniques for human identification.

1.3. SYNOPSIS OF THIS RESEARCH

Elliot and Collard assert that "there is a pressing need for bioarchaeologists and forensic anthropologists to develop more reliable methods for determining the ancestry of unidentified human remains" (2009:859). If ancestry determination is going to continue to be an integral component of biological profiles of unknown human remains, then the methodology must become standardized, consistent, and dependable.

The primary purpose of this research project is to apply three-dimensional landmark-based geometric morphometric methods to the process of ancestry determination using 3D computer-generated images of the crania of individuals whose geographic ancestral origins are of known provenance and derived from sub-Saharan Africa and Europe. The specific aim of this research is to explore whether there are statistically significant differences in cranial shape, size and/or form between sub-Saharan African and European individuals which could be used to assign ancestral group membership reliably and accurately from the cranium as a whole or from specific features and traits on the cranium. The application of three-dimensional geometric morphometrics to the process of ancestry identification, focused specifically on areas of

the skull that are used in traditional morphological analyses, should allow for greater precision, reliability, and accuracy. Furthermore, since the complex and irregular shape of the skull, including its curves and contours, can be retained using three-dimensional landmarks and semilandmarks, the geometric morphometric method should improve and enhance the overall statistical shape analysis of the cranial form.

2. LITERATURE REVIEW

This section provides a review of the traditional and current methods that are used in forensic anthropology and human identification to make assessments about ancestry from skeletal specimens, as well as a general background on the sites of the skull and craniofacial complex that display the most variation in terms of the ancestry of people originating from sub-Saharan Africa and Europe. Following this will be an introduction into geometric morphometric methods and how they have been applied within physical anthropology and human identification thus far, and a brief summary on the use of 3D images as a form of data for analysis in physical anthropology and human identification.

2.1. ANCESTRY AND HUMAN CRANIAL VARIATION

It is well recognized that people originating from different geographic regions look different. Ancestral differences, however, are not only manifested in someone's outward appearance but also in the skeleton. The skull, especially the area that makes up the facial region, is considered to be the most useful indicator of ancestry in the entire skeleton (Howells, 1973; White and Folkens, 2005; Klepinger, 2006; Randolph-Quinney et al., 2009).

Human cranial variation is complex and results from a number of factors most of which relate to natural forces caused by evolutionary responses to environmental adaptations (Molnar, 1998). Genetic isolation, heredity, and environmental influences, such as climate and diet, have all played roles in making us look the way we do (Molnar, 1998; Simmons and Haglund, 2005; Harvati and Weaver, 2006). Simmons and Haglund claim "human variation results from relative genetic isolation (endogamy) of populations for long periods of time, which accentuated particular characteristics in

each population" (2005:163). They also say that "while some variability is adaptively based, much of it is simply the result of the perpetuation of particular morphology due to breeding within a restricted area" (2005:163). With the advent of modern forms of transportation and with increased migration on a global scale, however, populations are composed more and more of individuals of mixed ancestry making it increasingly difficult to assign ancestry based on skeletal or cranial variation. Gene mixing may result in a varying outcome of the shape differences that are typically observed within the different groups (Simmons and Haglund, 2005).

A great deal of research has been conducted within physical anthropology with the intention of distinguishing between the major geographic populations from the skull and craniofacial region (Gill, 1998). Extensive research has been performed by studying the "relationship between biological characteristics of the living and their skeletons" (Sauer, 1992). Generally speaking, skulls can be grouped in relation to ancestry based on a varying degree of skeletal features on the skull or a combination of these features thereof. Forensic anthropologists commonly classify human cranial variation into one of four broad ancestral groups for identification purposes (Wilkinson, 2004; Randolph-Quinney et al., 2009). The labels commonly used in forensic anthropology for these ancestral groups are Caucasoid (including Europeans, Asians from the Indian Subcontinent, North and East Africans, Middle Easterners and Mediterraneans), Negroid (including West and Southern Africans), Mongoloid (including Southeast Asians, Inuit, and Native Americans) and Australoid (indigenous Australians, Pacific Islanders, Fijians and Papuans) (Wilkinson, 2004:84; Randolph-Quinney et al., 2009:14). There is some controversy associated with these classificatory terms and many anthropologists, working both inside and outside of physical anthropology, prefer to not use these terms citing them as outdated, empirically invalid,

and/or socially harmful (Black, 2000; Lieberman, 2001). Negroid and Mongoloid can be seen as offensive since they originated from and may be associated with the typological methods of racial classification and may have negative social connotations associated with them. Presently, however, the classificatory terms used to designate ancestry are not meant to propagate racist or typological views. Human variation is better understood than in the past and these labels are only meant to be used to distinguish between different geographic groups based on differing skeletal characteristics in a broad, informative, and unbiased way (Cartmill, 1998).

What follows is a summary of the characteristics and traits that have been routinely observed in cranial studies of individuals from the Caucasoid and Negroid ancestry groups, since the classification of these populations is the focus of this research. For the purposes of this research individuals with Negroid ancestry will be referred to as sub-Saharan Africans and those with Caucasoid ancestry as Europeans. These morphological traits are fundamental to the morphological ancestry assessment method. It is important to mention that although a forensic anthropologist may be able to assess ancestry from skeletal specimens fairly well using the skull, no specific trait or feature of the skull distinguishes one ancestral origin from another exclusively. One population may exhibit traits that are considered more frequently observed than in another population and vice versa. Overlap, in terms of ancestral skeletal traits, occurs between different groups and this must be taken into careful consideration when producing any kind of assessment based on ancestry (Byers, 2005). Since no single trait is a marker of ancestry, a suite of traits should always be examined, if possible, to suggest the ancestry of unknown individuals (Klepinger, 2006). However, certain skeletal traits are more frequently observed in some geographic groups than others (Klepinger, 2006).

In general, individuals with ancestral origins from Europe (Figure 1) tend to have a mesocephalic (long to rounded) skull shape with a sloping to upright frontal profile and often a pronounced external occipital protuberance (Wilkinson, 2004; Randolph-Quinney et al., 2009). Facial breadth is narrow to wide and facial height is medium to high (Randolph-Quinney et al., 2009). Nasal aperture form will be long and narrow and the inferior borders of the nasal margin will be sharp and deep (or silled) with a prominent nasal spine (Rhine, 1990; Gill, 1998; White and Folkens, 2005; Randolph-Quinney et al., 2009). The nasal root is highly arched and steepled with a depressed nasion (Gill, 1995). Interorbital distance is narrow and orbits are considered angular with a rhomboid-like form, sloping downwards and laterally (Rhine, 1990; Gill, 1998:300; Klepinger, 2006). Supraorbital margins are moderate to heavy and the cheek bones are sharp and receding (Wilkinson, 2004). The facial profile is characterized as orthognathic, which means that the facial region tends to look flat with little or no projection in the dental area along the midline (Bass, 2005). The palate is long and narrow, appearing parabolic in shape (Rhine, 1990; Gill 1998; Wilkinson 2004). The mastoid process is large and long (Wilkinson, 2004; Randolph-Quinney et al., 2009). The mandible is regarded as having an undulating lower border, pinched ascending ramus, slanted vertical ramus, straight gonial angle, and a prominent projection of the chin which may be bilobated (Rhine, 1990; Gill, 1998:300; Klepinger, 2006). Dentition tends to appear small and often crowded (Gill, 1995).

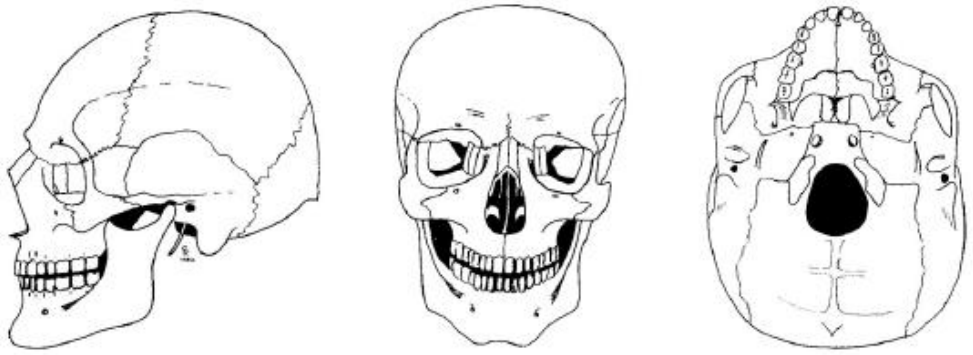

Figure 1 European type skull. (Figure obtained and adapted from Rhine, 1990).

In general, individuals of Sub-Saharan African ancestry (Figure 2) are characterized as having a dolichocephalic (long) skull shape with a sloping and rounded frontal profile and a post-bregmatic depression (Wilkinson, 2004; Klepinger, 2006; Randolph-Quinney et al., 2009). Facial breadth is narrow and facial height is low. Nasal aperture shape is wide and the inferior borders of the nasal margin are smooth or guttered, with little to no inferior definition (Rhine, 1990; White and Folkens, 2005; Randolph-Quinney et al., 2009). The nasal root is low and flat with a "Quonset hut" shape and a very dull or absent nasal spine (Randolph-Quinney et al., 2009:14; Gill, 1998:300; Gill, 1995; Rhine, 1990). Interorbital distance is wide and orbits appear rectangular (low and wide) in shape (Rhine, 1990). Supraorbital margins are somewhat undulating and are mild to moderate in prominence (Wilkinson, 2004). The cheek bones are receding with some lateral projection (Randolph-Quinney et al., 2009). The facial profile has pronounced prognathism, causing an anterior protrusion of the mouth region (Wilkinson, 2004). The palate is wide appearing hyperbolic to parabolic in form (Rhine, 1990; Gill, 1995; Gill 1998; Randolph-Quinney et al., 2009). The mandible has a narrow or pinched-looking ascending ramus with a straight inferior mandibular border and an oblique gonial flare (Rhine, 1990; Gill, 1995; Randolph-Quinney et al., 2009).

The appearance of the chin is blunt and vertical and dentition tends to appear large with little over-crowding as is often seen in individuals of European ancestry (Rhine, 1990).

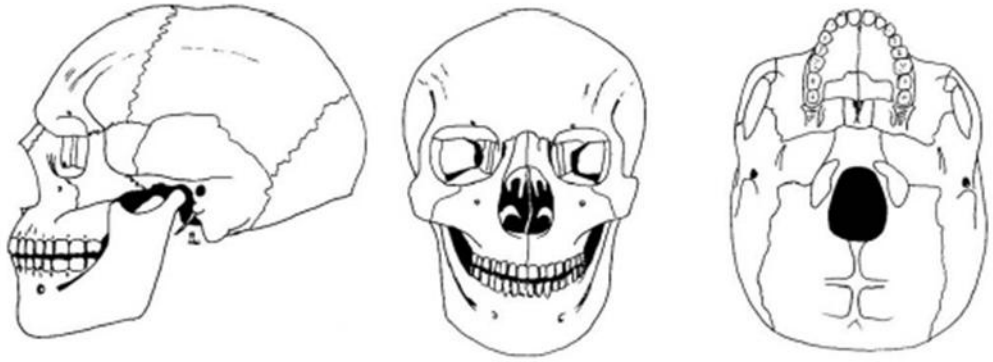

Figure 2 Sub-Saharan African type skull. (Figure obtained and adapted from Rhine, 1990).

The sites on the skull that are considered by anthropologists to show the most variation between individuals originating from different geographic areas occur within the nasal region, the supraorbital margin of the frontal bone, and the zygomatic and maxillary profiles (Gill and Gilbert, 1990; Gill, 1998). These regions are the predominant focus of this research project.

2.2. ANCESTRY ASSESSMENT APPROACHES

There are two approaches traditionally used for determining the ancestry of human remains using the skull: the morphological or non-metric method, and the anthropometrical or metric method. Both methods are based on the premise that there are variations in the shape and features of the skull between individuals originating from different geographic regions. The two methods are also used to predict sex from skeletal specimens using recognized sex-specific traits.

2.2.1. NON-METRIC APPROACH

Non- metric ancestry analyses are performed by the visual examination of the overall shape and features of the bones of the skull. Particular traits and features are explained by qualifying a bone's overall shape, or by stating the presence or absence of a feature, or by describing the trait along a continuum of possible variants that are considered to be reflective of human ancestral skeletal variation (Rhine, 1990; Hefner, 2009). The non-metric approach is widely used and is often the preferred approach for assessing ancestry from the skull (Gill, 1990; Byers, 2005; Hughes et al., 2011). This approach does not require equipment and is relatively quick and easy to carry out compared to metric analyses which can be time consuming and sometimes difficult to perform (Rhine, 1990). The non-metric approach allows for a number of skeletal shape differences, both subtle and not so subtle, to be assessed and described that may be quite difficult to analyze with other assessment techniques. For example, it is not easy, or in some cases even possible, to measure the prominence of the supraorbital ridge, or the degree of sharpness or guttering of the inferior nasal margin, or the presence or absence of shovel-shaped incisors, with traditional linear distance measurements used in metric ancestry assessments. These features and traits are often better described than measured. Also, in cases where the skull or cranium is fragmented, the non-metric approach may be the only viable approach since metric analyses often require multiple measurements to be applied on the entire cranium or skull to make an attempt at an accurate estimation of ancestry (Rhine, 1990).

The accuracy and reliability rates of the non-metric method for ancestry assessment from the skull have been tested by various researchers and have demonstrated mixed results. Some studies have shown that the non-metric approach can

be accurate and reliable in classifying groups, as well as repeatable between observers, if the descriptions of the various traits used are clearly defined and properly illustrated (Gualdi-Russo et al., 1999; Hefner, 2009; Hughes et al., 2011). Other studies, however, have suggested that the non-metric approach is problematic since the number of traits used for conducting analyses is not standardized across the field and the definitions of the appearance of the various traits are not adequately defined. It is argued that this results in significant differences between observers when interpreting the suggested appearance of a trait (Carpenter, 1976; Gualdi-Russon et al., 1999; L'Abbe et al., 2011). Some practitioners caution that the non-metric approach should be used only to supplement other more reliable and objective approaches; other practitioners recommend that this approach should not be used at all because of its serious shortcomings (Carpenter, 1976; L'Abbe at al., 2011). If clear definitions and illustrations of the skeletal features that differ between groups can be maintained across analyses perhaps the subjective variation in scoring different skeletal traits may be resolved or, at least, minimized. However, there is no consensus at present as to what constitutes an optimal set of traits that are most suitable for an approximation of ancestry using the skull (Hefner, 2009). Also, since no single trait defines a particular geographic ancestry category, and the traits observed are not fixed and are subject to individual variability within and between populations, perhaps it is not possible to come to a unified agreement within the field on exactly what traits should be utilized.

Despite all of these disagreements, what is at the heart of the problem with the non-metric approach is that the methodology used is intrinsically subjective in nature (Spradley and Jantz, 2011). There are not a lot of detailed guidelines or definitions by which to compare degrees of visible variation amongst practitioners and this can lead one observer, for example, to describe or score a particular trait as more or less

pronounced than another observer. Inter-observer variation is, therefore, a serious disadvantage with this method. The non-metric method relies heavily on the overall experience of the observer and puts those individuals that are new to the field at a disadvantage. Hefner argues that interpreting cranial variants in non-metric analyses is "a method that is as much an art as it is a science" (Hefner, 2009:985). There needs to be more standardization and testing of the methods used in non-metric analyses if this approach is going to continue to be used in biological profiling and as a potential source of expert evidence in a court of law (Hefner. 2009). So far, the non-metric approach to ancestry determination does not meet the guidelines established by the *Daubert* or *Mohan* ruling for determining whether evidence is scientific and therefore admissible in courts within the United States and Canada (Christensen, 2004; Rogers and Allard, 2004). The subjective nature of the non-metric approach to ancestry determination, which depends heavily on a practitioner's level of experience in handling particular skeletal specimens and features, does not demonstrate a scientific or objective technical disposition (Rogers and Allard, 2004). It is argued that the non-metric methods have "not been established as reliable or valid, they have not been subjected to appreciable peer review, and they have no known error rates" (Hefner, 2009:986).

2.2.2. METRIC APPROACH

Traditional metric methods for ancestry estimations involve osteometric measurements between standard points of interest, or landmarks, on the skull with possible subsequent statistical analyses. Metric analyses are considered more objective and more quantitatively precise than non-metric methods (Gill and Gilbert, 1990). Instructions about how and where to take measurements are detailed and more clearly defined in the literature, so there is less observer interpretation involved. In the 1960s,

Giles and Elliot created discriminant function formulae to classify skeletal specimens in terms of ancestry using linear distance measurements from a number of skulls representing different regional populations (Albanese and Saunders, 2006). This led to the development of computer programs, like Fordisc, which have been designed to estimate the ancestry and population origin of skeletal specimens through mathematical and statistical analyses of skull measurements (Ousley and Jantz, 2005; Elliott and Collard, 2009). Being able to apply statistical analyses to methods of ancestry determination gives the procedure a stronger scientific disposition, making it more effective in the courtroom "since it is an objective, quantitative method with some established confidence limits" (Gill and Gilbert, 1990:50)

One of the criticisms of the metric approach is that linear measurements leave out a lot of important shape information. For example, it is difficult to detect and assess the morphology of a region occurring between different landmarks on the skull when using conventional measuring devices like spreading and sliding callipers. Linear distance measurements do not pick up the contours or overall shape and complexity of the cranial form. Also, the size of the specimen is not taken into consideration. The size of a specimen may influence the way the overall shape is interpreted perhaps making it difficult to recognize and assess some shape differences if size difference is considerable between individuals. Rhine also asserts that many of the criticisms applied to non-metric methods are also relevant to metric analyses (Rhine, 1990). Rhine notes that a practitioner's experience is also a factor in metric analyses because there is always potential for mismeasurement since some landmarks may be difficult to locate and someone new to the field may be unfamiliar with the definition and placement of certain landmarks (Rhine, 1990). The metric approach is limited in cases in which there is poor preservation of remains. In general, metric analyses have to be carried out on

complete or mostly intact crania. On fragmented or incomplete crania, many of the required landmarks may be poorly located, or not locatable at all, seriously affecting the validity of the mathematical and statistical analyses of the measurements (Elliott and Collard, 2009).

The use of computer programs has become increasingly more common in metric methods of ancestry assessment in forensic anthropology. Programs such as Fordisc, created by Stephen Ousley and Richard Jantz, use standard cranial measurements applied to discriminant function analyses to assist in the determination of the ancestry, sex and stature of unknown human skeletal remains (Ousley and Jantz, 2005; Dirkmaat et. al, 2008:37). Particular areas of bones are measured from the unidentified human remains and these values are input into the computer program and compared to a database of skeletal measurements taken from the Forensic Data Bank and from William W. Howells' craniometric dataset, to ascribe group membership. The Forensic Data Bank is a large database that contains thousands of measurements of modern, extensively documented, and positively identified individuals from forensic cases within the United States. Howells' craniometric dataset consists of a large database of standard craniometric measurements that were taken by William W. Howells, an anthropologist, between 1965 and 1980 (Howells, 1996). Howells' craniometric data is freely available online and is comprised of measurements taken from over 2500 crania representing 28 regional populations which were derived both from historical and relatively modern skeletal collections. The population sets within the Forensic Data Bank database are based on commonly used folk taxonomy racial categories within the United States, such as Black, White, and Hispanic; those from the Howells' database are more population specific, representing groups from varying regional and temporal populations, such as the Norse, Ainu, Easter Islanders, and Egyptians. Using Fordisc to ascribe group

membership of ancestry and/or sex to unidentified human remains provides a means to "standardize the attribution of population affinity for forensic analysis and to provide a user-friendly method for ascribing "social race" to unknowns" (Williams et al., 2005).

Although Fordisc provides a standardized and objective technique for assessing ancestry, there are many studies that have commented on the overall inaccuracy of the program. Multiple studies, using crania and skulls with known sex and ancestry, have shown that Fordisc was unable to allocate properly the individuals into their appropriate ancestral or folk taxonomical population groups or to the nearest neighbouring ancestry population group (Kosiba 2000; Williams et al., 2005; Elliot and Collard, 2009). Furthermore, for skulls that were incomplete, which may often be the case in forensic and archaeological investigations, Fordisc results were unreliable (Elliot and Collard, 2009). Elliot and Collard conducted a study in which they evaluated the utility of Fordisc using the skulls of 200 individuals of known ancestry. They ran analyses both with and without the test specimen's source population included in the program's reference sample (Elliot and Collard, 2009) They found that Fordisc was able to classify more than 70% of individuals in analyses when the test specimen's source population was included in the reference sample and 56 measurements were employed but less than 40% of individuals were correctly classified when individuals were treated as unknowns and fewer measurements were utilized. Fordisc is only capable of identifying the ancestry of unknown individuals if the population of origin of the unknown is represented in the program's reference groups. The problem here is that a forensic practitioner using Fordisc will not know if the unknown individual's population group is one of the groups in Fordisc's database (Kosiba, 2000; Williams et al., 2005). If measurements from a particular skull are entered into Fordisc and the skull does not belong to one of the represented populations it may be statistically and inappropriately

'forced' into one of the program's ancestry groups. The program does statistically rate the classification result with a confidence or error level using posterior and typicality probabilities, however, various studies have indicated that the reference sample may be problematic since some populations represented are from archaeological and historical time periods and that attribution of ancestry should only be accepted if posterior and typicality probabilities are exceedingly high (Kosiba 2000; Williams et al., 2005; Elliot and Collard, 2009). This poses a serious problem because this program has been, and is still being, applied in forensic investigations with crania that may not be represented in the sample. The program's usefulness as a forensic tool to make ancestry assessments from skeletal remains has been shown to be highly problematic and even though this method may be more objective than other methods of ancestry assessment the program does not offer an adequate platform to make accurate and reliable estimations of ancestry.

2.3. GEOMETRIC MORPHOMETRICS

The study of shape and shape change are central topics to the discipline of physical anthropology as well as forensic anthropology. Physical anthropologists and forensic osteologists study the shape of the bones of the human skeleton to attempt to extract important biological and social information from it. The ability to combine both morphological and metrical approaches to the study of the shape and form of the human skeleton is important and geometric morphometrics may provide the means to do this. Geometric morphometric techniques provide the repeatability of metric analyses while still retaining all of the shape information of the form under study, thus, in a way, bridging the gap between traditional morphological and metrical techniques of analysis.

The study of morphometrics refers to the quantitative analysis of form, which consists of both size and shape, where mathematics and statistics play important roles in the analyses (Bookstein et al., 2004). Morphometric studies attempt to quantify the variations between different biological specimens statistically through the use of linear distance measurements, angles and/or ratios (Slice, 2005). Landmark-based geometric morphometrics, on the other hand, focuses on the two- or three-dimensional Cartesian coordinates of anatomical landmark points rather than on a collection of distance measurements, ratios and angles between certain anatomical loci (Slice, 2005). By using Cartesian coordinate points to study and quantify shape, all of the geometric information about an object can be retained throughout a study (Slice, 2005). In other words, the spatial arrangement between all of the landmark points from specimen to specimen is preserved throughout analyses (Mitteroecker and Gunz, 2009). Shape changes between groups of specimens are studied by examining the locations of the configurations of landmarks across all specimens relative to one another after important statistical shape analyses are performed. What especially sets this technique apart from traditional morphometric analyses is that the results obtained from geometric morphometrics studies can be easily understood as well as visualized by the researcher (Bookstein et al., 2004; Slice, 2005). With the application of geometric morphometric methods, shape changes that can be observed with the naked eye, as well as those changes that are more subtle and difficult to identify and measure, can be quantified statistically, expressed graphically, and represented visually (Kovarovic et al., 2011; Seetah et al., 2012).

Fundamental to geometric morphometrics methods are landmark coordinates. Landmarks are distinct anatomical loci that can be recognized as the same points in all specimens in the study (Zelditch et al., 2004). Within geometric morphometrics studies,

landmarks play a vital role and there are certain criteria for how they should be used and located. Zelditch et al. propose that landmarks should be "(1) homologous anatomical loci that (2) do not alter their typological positions relative to other landmarks, (3) provide adequate coverage of the morphology, (4) can be found repeatedly and reliably, and (5) lie within the same plane" (Zelditch et al., 2004:24). Zelditch et al. recommend that landmarks should be chosen to answer the questions of the particular study being carried out, but they should also be chosen on a broader scale so as not to limit the potential discovery of new shape information (2004).

Bookstein et al. classified and ranked landmarks in terms of the ease of identifying their locations on biological structures. Landmarks are commonly classified into three types: Type I, Type II and Type III (Bookstein et al., 2004). Type I landmarks are those which are homologous, meaning that they have the same relative position across all specimens. They have a strong locally defined histological location making them easily locatable and repeatable among specimens and researchers. Examples of Type I landmarks on the skull include an intersection between three sutures (such as bregma or asterion) or the location of small foramina (Hallgrimsson et al., 2007). Type II landmarks are classified as points whose homology from specimen to specimen are maintained by geometric or shape evidence and not based on locally defined biology. Points of greatest curvature along a bony edge, such as the landmark jugale, or the tips of the ends of bony projections, such as mastoidale, or the intersection of sutures with edges, such as frontomalare orbitale along the eye orbit, would all be classified as type II landmarks (Hallgrimsson et al., 2007). Type III landmarks are more difficult to locate and, hence, are considered more deficient. They are usually extremal points that lack proper definition as to their exact location across all specimens. Examples of Type III landmarks include ectoconchion and euryon, which are defined as points with extreme

curvature or the points furthest away from a particular structure (Hallgrimsson et al., 2007).

Semilandmarks can also be used in geometric morphometrics studies. In two-dimensions, semilandmarks are a series of points along an outline and in three-dimensional studies they are points along curves or surfaces. They would be classified as the most deficient of all landmarks since their exact location on a biological specimen can rarely be properly named (Mitteroecker and Gunz, 2009). Semilandmarks are used to include information about homologous features on specimens that lie between landmark points (Gunz et al., 2005). Although they lack exact positional information, semilandmarks can be extremely useful and important because they can provide geometric shape information about a biological form that cannot be obtained from conventional landmarks.

Landmarks are generally captured in a variety of ways with a range of different devices. Two-dimensional landmarks are usually acquired using a digitizing tablet or by measuring and landmarking digital photographs on a computer screen (Mitteroecker and Gunz, 2009). Three-dimensional landmark coordinates can be obtained directly using coordinate digitizing devices such as the Polhemus Patriot or Microscribe, or they can be digitized from virtual objects (created from surface or volumetric scans of real objects) visualized in three-dimensions on a computer screen using programs such as Landmark, EVAN, and Checkpoint (Mitteroecker and Gunz, 2009).

In traditional morphometrics, form is studied quantitatively. Form is understood to consist of both shape and size (Richtsmeier et al., 2002). In some geometric morphometrics applications, size is removed as a factor allowing pure shape to be studied. Shape, in geometric morphometrics, is defined as "all the geometric

information that remains when location, scale and rotational effects are filtered out from an object" (Kendall, 1977). Being able to distinguish between size and shape in morphological analyses is an important advantage since the size of a specimen can have an effect on the measurements taken and subsequently affect the results obtained. Traditional metric analyses utilized in forensic anthropology and human identification generally do not single out shape as a distinct variable from size and this can impact the interpretations made about morphological variation.

There are several different methods and models available to analyze landmark coordinates. These include Procrustes superimposition techniques, Euclidean distance matrix analysis (EDMA), finite element scaling analysis (FESA), and thin-plate spline analysis (TPS) (Adams et al., 2002). The choice of method and model is dependent on a number of factors, including the nature of the structure being analyzed, the choice of the types of landmarks being used, and the particular research questions under investigation. The most commonly used geometric morphometrics method to analyze the positions between landmark points is the Generalized Procrustes Analysis (GPA), which is also known as Procrustes superimposition (Viscosi and Cardini, 2011). GPA is the most theoretically developed method of geometric morphometric landmark-based analyses and is often the preferred or "default" method utilized among biologists and morphometricians (Zelditch et al., 2004; Viscosi and Cardini, 2011; Cardini, 2012). A thorough explanation of why the Procrustes method is preferred is available in the textbook Geometric Morphometrics for Biologists by Zelditch et al. (2004). Two of the main strengths of the method are its consistency with the mathematical theory of shape and its advantageous statistical properties (Cardini, 2012). The Procrustes method standardizes the landmark configurations by using mathematical algorithms and adjusts the raw landmark data so that all the specimens in the study are superimposed, or

aligned, into a common coordinate system. This allows for the data to be collectively analyzed and compared since the landmark points are placed in similar orientation and are adjusted for size differences. All the information about the landmark configurations that is not related to shape is removed (Rohlf, 2003). GPA translates the data so that all data is centred together around a common origin (or centroid), scales all of the data to a common unit centroid size, and optimally rotates all of the data to the same orientation (O'Higgins and Strand- Viðarsdóttir, 1999; Slice, 2005; Viscosi and Cardini, 2011). When GPA is applied to raw landmark data an average mean shape (or centroid) of each landmark configuration is constructed and all of the landmark configurations are translated, or re-distributed, around the new average shape, which is centred at the origin (Weber and Bookstein, 2011). The size of each specimen is estimated by measuring the scatter of the landmarks around the mean and the landmark coordinates are scaled to the same unit centroid size by minimizing the sum of squared distances between each landmark and the corresponding centroid points (Adams et al., 2004; Cardini, 2012; Slice, 2007). Landmark coordinates are then rotated so that the Procrustes distances between the specimens are reduced, again, by minimizing the squared distances between landmark configurations and the centroid (Adams et al., 2004). Size is removed from the shape data and a separate size variable is created. Since size is transformed and stored as a separate variable, the researcher has the options to explore shape variables independent from size, size variables independent from shape or, if desired, to use size in conjunction with shape for studies of form and allometry. After superimposition of the landmark data, differences in the landmark coordinate locations reflect differences in the shapes of the specimens (Slice, 2007). Landmark coordinates can then be statistically analyzed as shape variables using ordination methods, such as principal component analysis, or multivariate statistical analyses, such

as discriminant function and canonical variates analyses, to identify and quantify group differences (Slice, 2007).

There are some limitations associated with Procrustes-based geometric morphometric methods. One of the main limitations that is often discussed in the literature is the 'Pinocchio effect' (Zelditch et al., 2004; von Cramon-Taubadel et al., 2007; Viscosi and Cardini, 2011; Cardini, 2012). After Procrustes superimposition is performed on raw data, the positions of the landmarks represent shape differences as a whole but the method cannot necessarily convey the amount of shape variation occurring at individual landmarks between individuals (Viscosi and Cardini, 2011). If significant variation between landmark configurations is limited to only one or a few landmarks within the configuration, then the variation between these landmarks may be smeared out across all of the landmarks used in the study (Zelditch et al, 2004; von Cramon-Taubadel et al., 2007). The least-squares criterion that is used to superimpose landmark configurations may spread the shape displacements of the most variable landmarks across the entire configuration of landmarked specimens (Zelditch et al., 2004). The fairy tale of Pinocchio and his growing nose illustrates this effect well (Viscosi and Cardini, 2011; Cardini, 2012). If Pinocchio's head were to be landmarked both before and after the length of his nose changed from lying and these landmark configurations were subjected to Procrustes superimposition, the variation observed in the landmark configurations would suggest that Pinocchio's whole head changed shape along with the length of his nose even though his nose is the only feature that is altered (Cardini, 2012). Figure 3 illustrates the 'Pinocchio effect' of a Procrustes superimposition. Procrustes superimposition can minimize the variation in shape between different individuals by distributing the variation between landmarks evenly over the entire configuration. This may, however, misleadingly decrease the variation

occurring at different landmarks and, possibly, generate inconsistent estimators of mean form and shape (Lele, 1993). However, studies have shown that the Procrustes method is more accurate at estimating the true mean shape of a configuration of landmarks than other geometric morphometric methods (Rohlf, 2003).

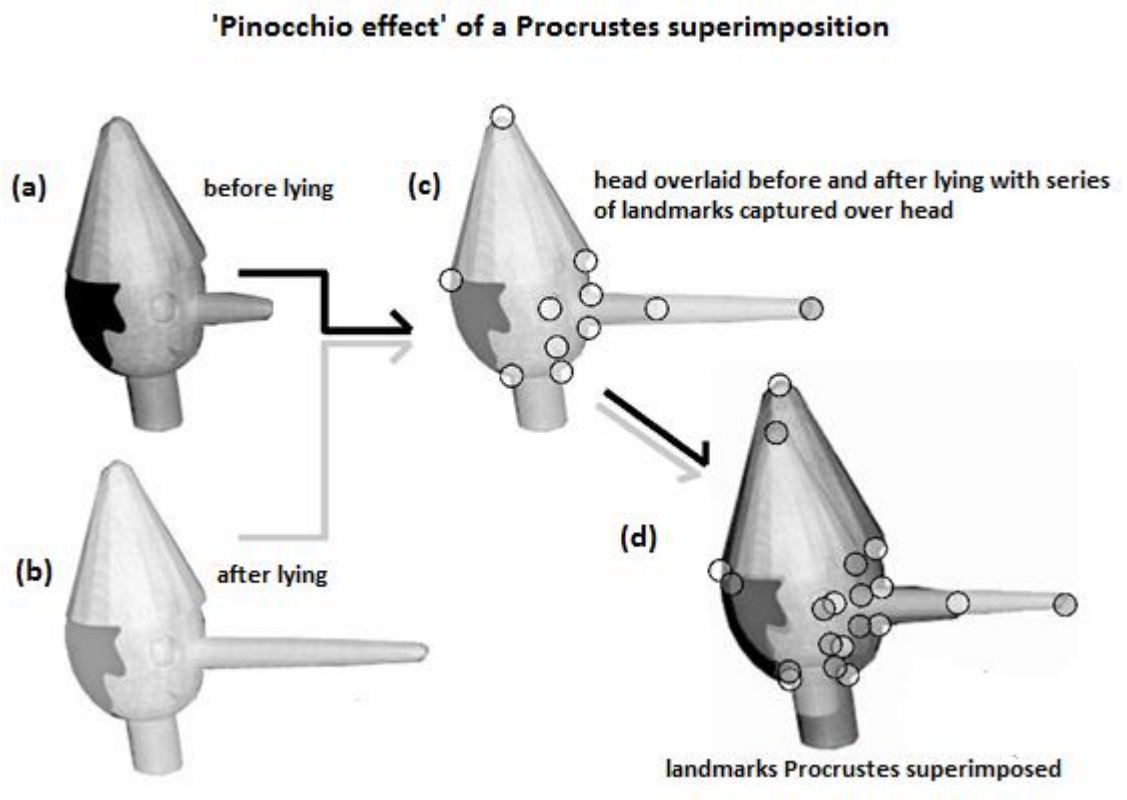

Figure 3 Visualization of the Pinocchio effect from Procrustes superimposition.

(a) Pinocchio's head before lying, (b) his head after lying. Pinocchio's head overlaid (c) before and after lying with a selection of landmark points over his entire head. The only shape difference that can be observed is the length of his nose. However, after Procrustes superimposition (GPA) (d) shape differences are observed over his entire head instead and are not confined to his nose, where the only actual shape difference is found. (Figure obtained and adapted from Cardini, 2012, In press)

Euclidean Distance Matrix Analysis, or EDMA, is another form of landmark-based geometric morphometric analysis. EDMA does not register sets of landmarks so it does not rely on superimposition methods or a common coordinate system (Lele, 1991;

Lele and Richtsmeier, 1991a; Lele and Richtsmeier, 1991b; Cole and Richtsmeier, 1998). EDMA uses a form matrix and evaluates shape differences between two specimens by looking at all of the possible Euclidean distances (linear distances) between homologous landmarks. A form matrix is constructed from the Euclidean distances from each specimen and the matrices are compared to one another to identify how the linear distances may differ between specimens (Ferrario et al., 1993). EDMA provides a method to identify and single out influential shape landmarks based on the linear distances that are the most variable between specimens (Ferrario et al., 1993). The comparison of all possible inter-landmark distances as a selection of ratios prevents the need to register landmark configurations (O'Higgins and Jones, 1998). There are proponents for and against the use of EDMA *vis à vis* superimposition techniques. EDMA does not rely on an arbitrary registration system but the interpretation and visualization of the method's results can be more difficult to perform and express (O'Higgins and Jones, 1998). Linear distances between landmark positions do not remove the issue of size from the data so mathematical algorithms to adjust for size differences must be utilized (Cole and Richtsmeier, 1998).

2.3.1. Geometric morphometrics in anthropology and human identification

The application of modern morphometric techniques to anthropological analyses is a relatively new and exciting aspect of research, especially within the field of human evolution and forensic anthropology. Over the last ten years geometric morphometric investigations in physical anthropology have greatly increased particularly in palaeoanthropology in studies of morphological change due to human evolution (Yaroch, 1996; Humphrey et al., 1999; O'Higgins, 2000; Hennessy and Stringer, 2002; Martinez-Abadias et al., 2006; Bruner and Manzi, 2007). In the past, these methods

were predominantly used in the biological and biomedical sciences but Bookstein and colleagues claim that the field of anthropology may in fact be taking over the use of geometric morphometrics (2004). In more recent years the fields of forensic anthropology and human identification have begun to adopt the methods of geometric morphometrics for studies addressing human variability and biological identity (Buck and Viðarsdóttir, 2004; Oettlé et al., 2005; Bytheway and Ross, 2010; Gonzales et al., 2011; Nicholson and Harvati, 2006; Pretorius et al., 2006; Braga and Treil, 2007; Franklin et al., 2007; Kimmerle et al., 2008; Husmann and Samson, 2011; Sholts et al., 2011b). Most of the studies concerning geometric morphometrics in forensic anthropology focus on studies of sexual dimorphism using the cranium, mandible or pelvic bones, although there are some studies in the literature involving analyses on the assessment of age at death as well as ancestry.

Oettlé et al., (2005) examined human mandibular rami using two-dimensional geometric morphometrics to determine whether significant differences exist between the sexes using this feature alone. They found that the mandibular ramus showed considerable overlap between the sexes and that there is insufficient dimorphism in this trait alone to estimate the sex of unknowns. Pretorius et al., (2006) utilized two-dimensional geometric morphometric methods to examine the shape of the greater sciatic notch, mandibular ramus flexure, and shape of the orbits to identify sexually dimorphic characteristics from photographs. As they anticipated, they found that the shape of the greater sciatic notch was the most sexually dimorphic of all the morphological characteristics they studied. However, they were surprised to find that the shape of the orbits differentiated between the sexes more than the shape of the ramus flexure. Generally, orbital shape is considered a less sexually dimorphic characteristic than the shape of the mandibular ramus which has been considered a

reliable morphological trait for sex identification (Loth and Henneberg, 1996). Another study by Gonzalez et al. (2011) looked at sexual dimorphism in crania using geometric morphometrics with two-dimensional landmarks and semilandmarks focused specifically on structures that are considered sexually dimorphic. They found that the percentage of correct sex classification from cranial traits, such as the glabellar region, mastoid process and frontal and zygomatic processes, was low, suggesting that in general there is a low level of sexual dimorphism in the shape of the cranium from these features. The cranium, however, is considered to be the second most useful indicator of sex from the skeleton since many studies have reported high accuracy rates using metric techniques and morphological features of the cranium to determine sex (Galdames et al., 2008). However, when they used both shape and size variables together their percentage of correct sex classification increased markedly, "indicating that the traits analysed display marked sex differences related to the larger size and more robust features of males" (Gonzalez et al., 2011:82).

There are few studies in the literature that specifically relate geometric morphometrics to ancestry determination. Ross et al., (1999) conducted a preliminary study in which they used the "new morphometry" to attempt to allocate 19 American Black and 19 American White crania into their respective ancestry groups. They used fourteen cranial landmarks and measured them in sequences of three using calipers and then subjected these interlandmark distances to mathematical analyses to extrapolate *x* and *y* landmark coordinates. After superimposition, they created a wireframe of the mean configurations of the landmarks for Blacks and Whites using a computer program. They overlaid the mean configurations of the landmarks from both groups and were able visually to compare the overall change in position of the landmarks between the groups. They were able to classify 78.9% of Blacks and 88.9% of Whites correctly

using particular landmark and semilandmark coordinates. They compared these results to traditional ancestry discriminant analyses, using cranial distance measurements, and found that the two types of analyses produced very similar results. They concluded that although the outcome of their analysis resulted in similar correct classifications for both techniques, geometric morphometrics methods were more promising because they could present information easily about the specific location of important morphological variation relating to ancestry and have the added benefit of displaying morphologically informative shape diagrams that could not be achieved easily through traditional methods of ancestry determination using discriminant analysis.

Henessy and Stringer (2002) provide a detailed examination of the practicality of using geometric morphometrics to study the regional variation observed in human craniofacial form. They used a three-dimensional laser scanner to digitize the surfaces of skulls of known ancestry, including Inuit, African, Australian and British individuals. The three-dimensional coordinates of nine craniofacial landmarks were extracted from the laser scans and analyzed. They found that geometric morphometric methods were able to characterize the overall face shape and variation within each sample accurately as well as distinguish the chief differences between the samples. Another recent study involving ancestry and geometric morphometrics was carried out by Husmann and Samson (2011). Their study involved determining the race and sex of individuals using only the orbital aperture. By applying 2-dimensional geometric morphometric principles to study the shape of the orbital aperture they found that there was very little variation in the shape of the orbits that could be attributed to race or sex differences. They concluded that the shape of the orbital aperture alone should not be used to determine the sex or race of unknowns.

A relatively new program called 3D-ID has been developed by Ann H. Ross and Dennis E. Slice to estimate the sex and ancestry of unknown human cranial remains using the principles of landmark-based three dimensional geometric morphometric shape analysis (Slice and Ross, 2009). The user can capture and plug in up to 34 three-dimensional cranial landmark coordinates and the software compares the landmark coordinate configurations to a database of over 1000 individuals of known sex and ancestry. The program attempts to classify the cranium of interest by comparing the shape of the cranium to the shape information obtained from the reference sample through optimized and landmark-specific classification functions (Slice and Ross, 2009). Classification results are based on Mahalanobis squared distance from the cranium of interest to each available reference group's mean shape. The program also provides diagnostic statistics such as posterior probabilities of membership into the different reference groups as well as the typicality probability for how the cranium of interest compares with other specimens within the reference sample (Slice and Ross, 2009).

Although the program was released in 2010 there is currently little information in the relevant literature about 3D-ID. How well it performs, how accurate it is, and how widely it is used by practitioners in forensic investigations is not known. The choice of landmarks used in 3D-ID represent strictly Type I and Type II landmarks, which were shown to be more repeatable in an intra- and inter-observer landmark precision study conducted during the initial phase of the making of the program (Ross et al., 2010). However, many standard anatomical landmarks located around key areas of the craniofacial complex that show distinct ancestral variation, that would be considered repeatable since they are Type I and Type II landmarks, are not utilized (Lockeyer, 2010). For example, no landmarks on the nasal aperture are utilized, and the landmark

infranasion on the nasal root is missed, as well as landmarks along the dental arcade, which could detect differences in palate shape and prognathism. A former student's undergraduate thesis at the University of Dundee at the Centre for Anatomy and Human Identification attempted, in part, to test the accuracy of 3D-ID for ancestry classification (Lockeyer, 2010). Lockeyer found that 3D-ID did not fare well in classifying crania by ancestral origin accurately. Classification rates using 3D-ID for crania of known ancestry from this study achieved a correct classification 60% of the time. He concluded that the set of landmarks used in 3D-ID is not sufficient to summarize the areas of ancestral variation observed in the skull. Many of the 34 anatomical landmarks used by the program do not focus on areas of the skull that frequently show the most variation between ancestral groups.

Although some of the aforementioned studies have proven unsuccessful in the determination of sex or ancestry from skeletal remains using geometric morphometrics they still provided a significant step forward in quantifying and addressing skeletal shape differences among and between individuals. In terms of forensic anthropology and osteological examinations, geometric morphometrics is an extremely powerful tool for the researcher because it provides a quantitative and, hence, more objective means to analyze shape variation. There are various problems and constraints with the traditional quantitative and qualitative techniques applied to study the shapes and features of bones. Geometric morphometric techniques may provide a more specific and detailed analysis of human osteological variation within the study of forensic anthropology. Geometric morphometric applications are not necessarily intended to replace existing identification methods but to enhance and extend them, and, more generally, increase knowledge of the shape and form of the human skeleton.

Geometric morphometrics methods applied specifically to human ancestry determination using the craniofacial skeleton is a modestly explored area of research. It has the potential to be an important and cutting edge method for determining ancestry from human remains in a forensic context. Geometric morphometrics research has the potential to greatly increase knowledge of the shape and size relationship of craniofacial features as well as possibly define new areas of the skull as important markers of ancestral differentiation (Kimmerle et al., 2008). Since ancestry determination can be challenging, applying geometric morphometrics to study this aspect of the biological profile has the potential to simplify, standardize, quantify, and objectify ancestry research.

2.4. USE OF 3D IMAGERY IN ANTHROPOLOGY

The use of computed tomography (CT) scans and surface scans as representations of skeletal structures has become increasingly common in anthropological studies. The use of these scans as potential sources of data in anthropological investigations has a number of benefits over analyses that are performed directly on osteological material. 3D images provide a non-destructive and non-invasive environment to analyze materials that may be fragile and difficult to handle (Ramsthaler et al., 2010). Osteological remains from archaeological and forensic investigations may be fragmented and delicate and the use of 3D imaging techniques provides a means to analyze these types of materials with less handling of the actual bone. CT data also has the added benefit of displaying external as well as internal structures, such as the endocranium, sinus, and tooth roots, which may be hidden or difficult to access when using traditional methods of analysis (Weber et al., 2001). Specifically, within the field of human identification, CT scans of osteological materials may avoid the need for maceration procedures since

the bone can be visualized without the need to de-flesh it (Robinson et al., 2008; Ramsthaler et al., 2010). This may accelerate identification of the remains of the unknown individual since maceration can be a difficult, labour intensive and lengthy process. 3D images also provide the possibility to limit the amount of contact made with contaminated remains that may pose health and safety concerns (Robinson et al., 2008; Ramsthaler et al., 2010). Another major benefit of 3D imaging is that the object scanned can be permanently archived and can be rapidly re-distributed electronically to others on an international scale (Robinson et al., 2008; Ramsthaler et al., 2010). Finally, 3D surface and CT scans are well-matched to morphometric analyses since many programs are now readily available to visualize and mathematically analyze such forms of data.

Although there has been extensive research conducted showing the validity and reliability of 3D images for morphological and metric anatomical analyses, there are some concerns involved when using this type of data for scientific research purposes (Williams and Richtsmeier, 2003; Robinson et al., 2008; Ramsthaler et al., 2010; Sholts et al., 2011a). Three sources of potential errors in imaging systems have been identified by Richtsmeier and colleagues: (1) the quality of the scanning device used; (2) the ability of observers to reliably and precisely locate landmarks on 3D images; and (3) the ability of the observer to use the applied software correctly (Richtsmeier et al., 1995). Since surface scan devices, such as laser scanners, use different imaging technology than CT imaging, surface scans have the inherent limitation of lower resolution than CT scans. This may result in difficulty visualizing certain features on 3D images of surface scans, such as sutures, since they may not always display a sharp visual contrast on specimens especially if the path of the suture is faint or obliterated or the specimen displays signs of deterioration or discolouration (Sholts et al., 2011a). This can hold true for 3D images of CT scans also. However, in general, CT data is more effective at

capturing the true form and fine detail of an object than surface scan data (Williams and Richtsmeier, 2003).

3. MATERIALS AND METHODS

To capture the morphology of the crania used in this research, the Cartesian coordinates of sixty-eight anatomical points were measured that are commonly employed in morphometrics. These points have a clear one to one correspondence across individuals and are called landmarks. A series of semilandmarks, which are points on curves or surfaces where 'traditional landmarks' are not available, were also used to obtain detailed shape information on regions whose boundaries are defined by landmarks. The landmarks utilized are listed and defined in Table 1. In terms of their degree of homology and biological correspondence across specimens, five (or about 7%) were Type I landmarks (points in which three structures meet or have a strong locally defined histological location), fifty-six (or about 82%) were Type II landmarks (points that are located by geometrically defined correspondence across specimens), and seven (or about 10%) were Type III landmarks (points that are mostly based on extremities of features and structures) (Zelditch et al., 2004; Hallgrimsson et al., 2007). Three different areas on the cranium were chosen to be outlined by semilandmarks: the contour of the cranial vault traced along the midplane, the outer rims of the orbits (both left and right), and the outer edge of the lower nasal aperture. All of the areas marked by semilandmarks are considered important features in traditional ancestry assessments since they are considered ancestrally dimorphic. A complete description of the curves and outlines marked by semilandmarks can be found in Table 2.

Landmarks and semilandmarks were chosen primarily along areas of the craniofacial region that reflect the variation observed between different geographic populations. Such areas include the nasal region, consisting of the nasal root and the shape and width of the lower nasal aperture; orbital shape; orientation of the zygomatic

bones; shape of the maxillary profile; palatal shape; mastoid length; frontal profile shape; and prominence of the supraorbital ridges. Landmarks and regions marked by semilandmarks were chosen to be easily identifiable on crania and repeatable among researchers as well as to account for the variations that are routinely observed between ancestry groups in traditional non-metric and metric analyses.

Table 1 Cranial landmarks used in study and definitions

Definitions and abbreviations adapted from Martin (1928), Martin and Saller (1957) and updated by Moore-Jansen et al. (1994) and White and Folkens (1991).

Order placed	Landmark name and abbreviation	Type	Definition
Midline Landmarks			
1	Glabella (g)	III	Most prominent point on the frontal bone occurring between the supraorbital ridges in the midsagittal plane and usually above the nasal root
2	Nasion (n)	I	Point on the midsagittal plane where the frontonasal sutures meet
3	Rhinion (rhi)	II	Most anterior point at which the nasal bones meet
4	Subspinale (ss)	II	Deepest midline point on the maxillae between the nasal spine and prosthion
5	Prosthion (pr)	II	Most anterior midline point on the alveolar process of the maxillae between the central incisors
6	Bregma (br)	I	Point where the coronal and sagittal sutures meet
7	Lambda (l)	I	Point where the sagittal and lambdoidal sutures meet
8	Opisthion (o)	II	Midline point at the posterior rim of the foramen magnum
9	Basion (ba)	II	Midline point at the anterior rim of the foramen magnum
10	Posterior edge of incisive canal (ic)	II	Most posterior edge of the incisive canal on the hard palate
Bilateral Landmarks			
11 and 12	Infranasion (inf)	I	Most supero-lateral point on frontonasal suture
13 and 14	Naso-maxillary (nm)	II	Most infero-lateral point on naso-maxillary suture
15 and 16	Alare (al)	III	Most lateral point on the anterior margin of the nasal aperture
17 and 18	Superior zygotemporale (zts)	II	Most superior point on the zygotemporale suture
19 and 20	Inferior zygotemporale (zti)	II	Most inferior point on the zygotemporale suture
21 and 22	Zygomaxillare (zm)	II	Most inferior point on the zygmaticomaxillary suture

23 and 24	Zygoorbitale (zyo)	II	Point where the zygomaticomaxillary suture meets the orbital rim
25 and 26	Jugale (ju)	II	Deepest point in the notch between the temporal and frontal processes of the zygomatic
27 and 28	Frontomalare orbitale (fmo)	II	Point where the frontozygomatic suture meets the inner rim of the orbit
29 and 30	Frontomalare temporale (fmt)	II	Most lateral point where the frontozygomatic suture meets the temporal line
31 and 32	Frontotemporale (ft)	II	Most anteromedial point on the temporal line on the frontal
33 and 34	Ectoconchion (ect)	III	Most lateral point on the orbital rim
35 and 36	Max. orbital height (obh)	III	Most superior point on the orbital rim
37 and 38	Orbitale (orb)	III	Most inferior point on the orbital rim
39 and 40	Porion (por)	II	Most superior point on the border of the external auditory meatus
41 and 42	Auriculare (aur)	II	Point very close to and above porion which is situated at the root of the zygomatic process
43 and 44	Mastoidale (mast)	II	Most inferior point at the tip of the mastoid process
45 and 46	Asterion (ast)	I	Ectocranial point where the lambdoial, parietomastoid, and occipitomastoid sutures meet
47 and 48	Lateral foramen magnum (lfm)	II	Most lateral point on the border of foramen magnum
49 and 50	Endomalare (enm)	II	Lingual point on the alveolar margin of the maxilla that is located opposite of the centre of the second molar
51 and 52	Ectomalare (ecm)	II	Most lateral buccal point on the alveolar margin of the maxilla usually at the position of the second molar
53 and 59	Buccal 1st molar alveolus (m1)	II	Most buccal point on the alveolar margin of the maxilla between the first and second molars
54 and 60	Buccal 2nd premolar alveolus (p2)	II	Most buccal point on the alveolar margin of the maxilla between the second premolar and the first molar
55 and 61	Buccal 1st premolar alveolus (p1)	II	Most buccal point on the alveolar margin of the maxilla at the first premolar and second premolar
56 and 62	Buccal canine alveolus (c1)	II	Most buccal point on the alveolar margin of the maxilla between the canine and first premolar
57 and 63	Buccal 2nd incisor alveolus (i2)	II	Most buccal point on the alveolar margin of the maxilla between the second incisor and canine
58 and 64	Buccal 1st incisor alveolus (i1)	II	Most buccal point on the alveolar margin of the maxilla between first and second incisor
65 and 66	Most prom. Point on side of glabella (lg)	III	Most prominent point to the left of the glabellar region
67 and 68	Most prom. Point on superciliary ridge (supra)	III	Most superior point of projection on the supraorbital ridge

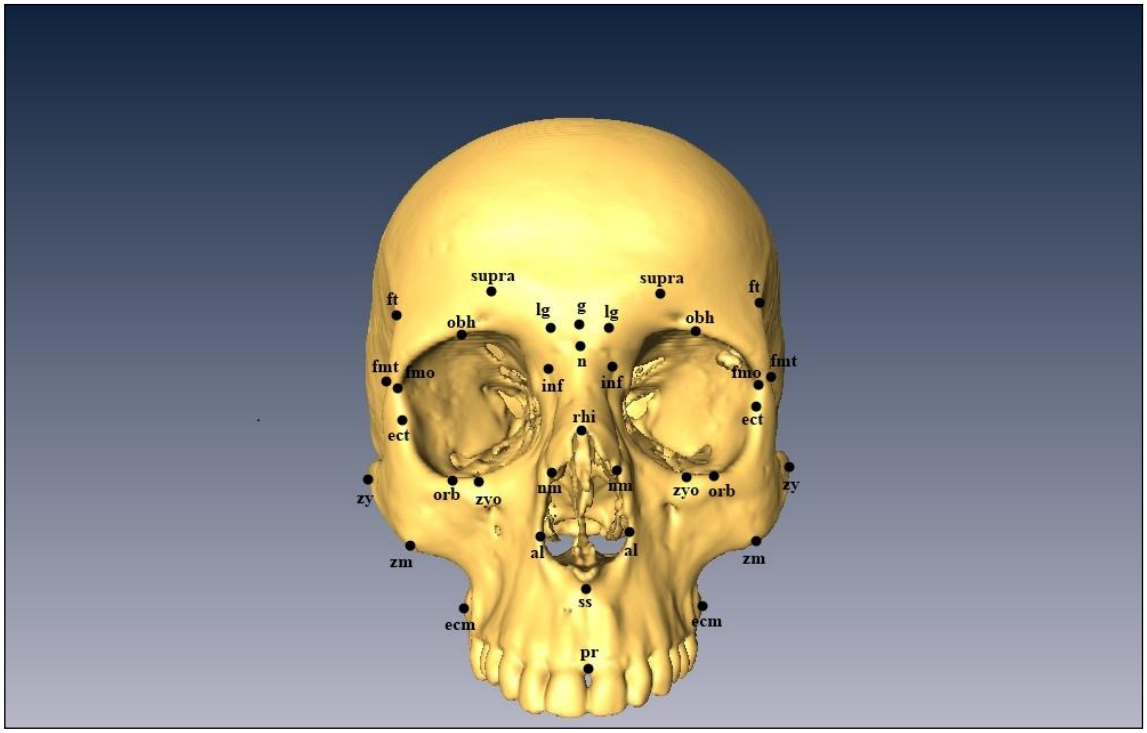

Figure 4 Frontal aspect of cranium with landmarks used in the study

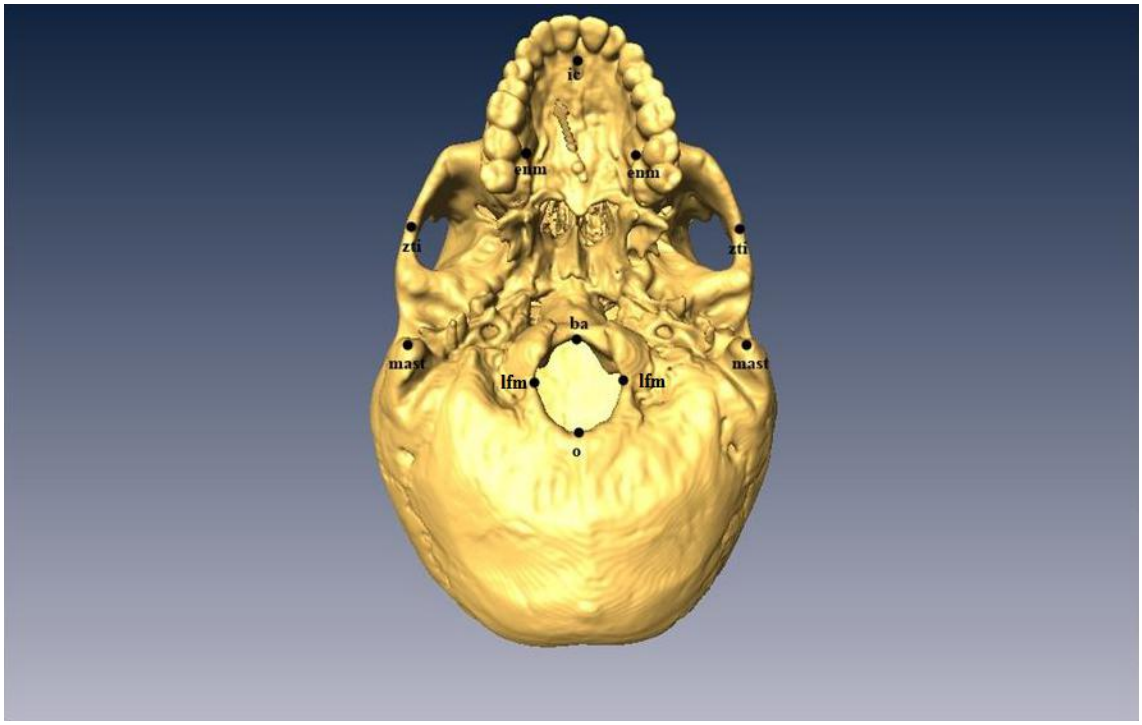

Figure 5 Inferior aspect of cranium with landmarks used in the study

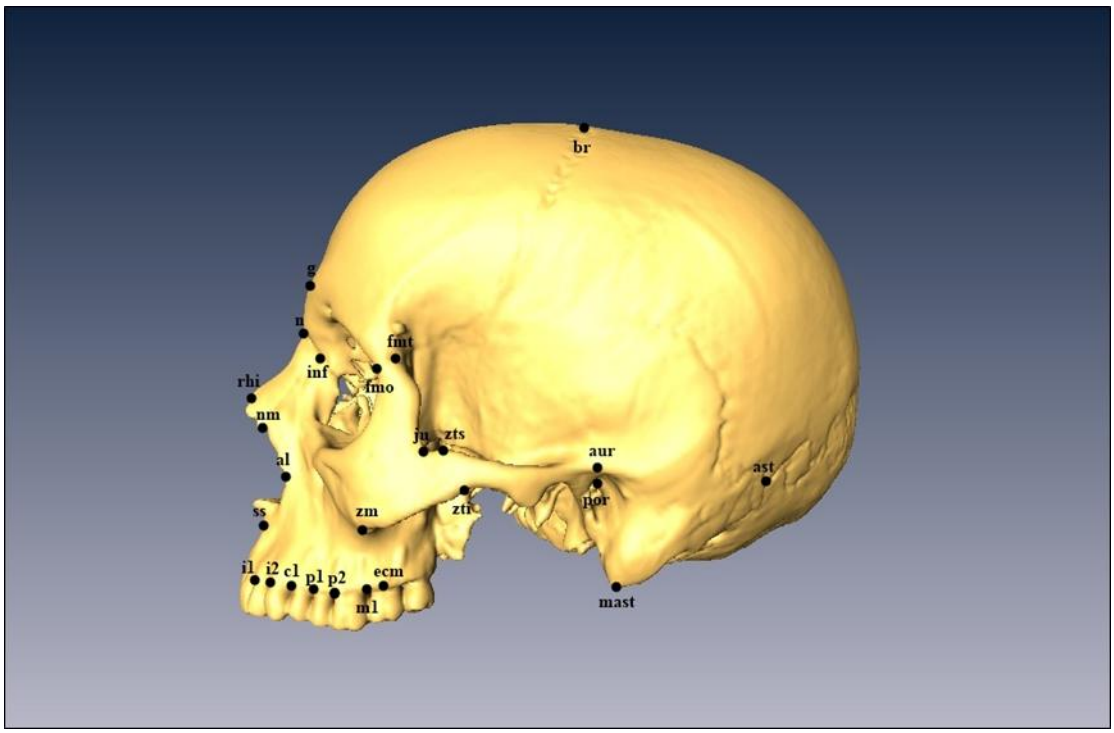

Figure 6 Lateral aspect of cranium with landmarks used in the study

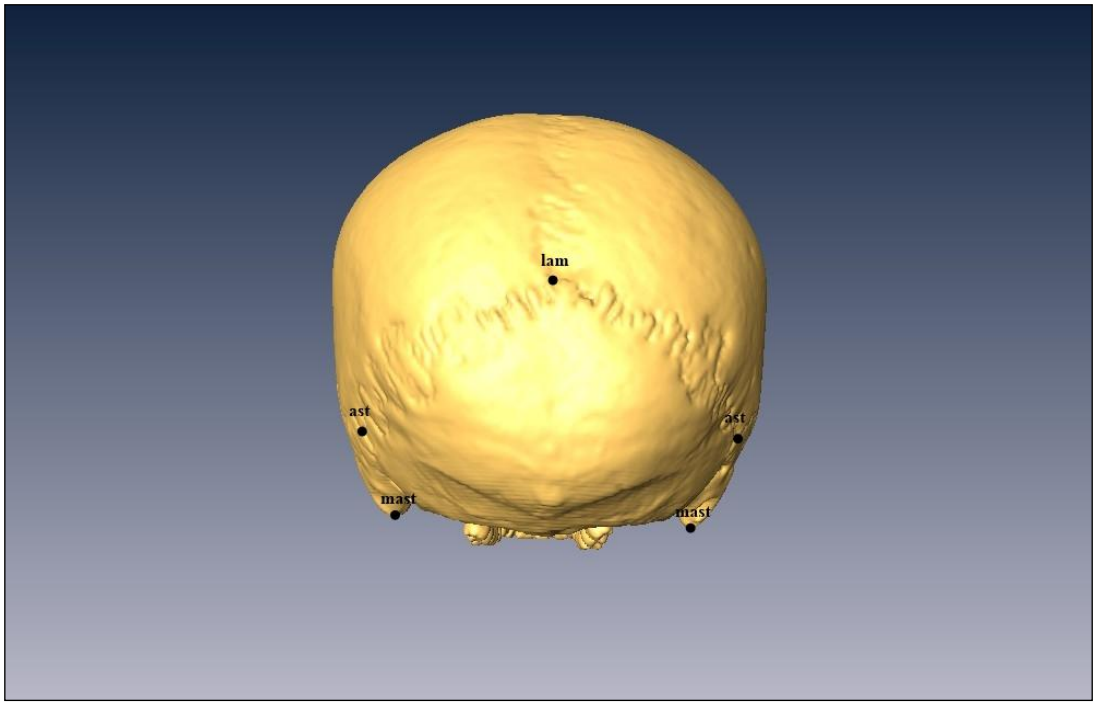

Figure 7 Posterior aspect of cranium with landmarks used in the study

Table 2 Description of semilandmarked areas used in study

Region and No. of points used	Description of curve	Start point and endpoint of curve or outline
Vault (30)	Trace of the contour of the cranial vault on the midplane	Rhinion to bregma; Bregma to lambda; Lambda to opisthion
Orbit (20) (both left and right)	Trace of the outer rim of the orbit	Frontomalare orbitale to frontomalare orbitale
Lower nasal aperture (10)	Trace of the outer edge of the lower nasal aperture	left alare to subspinale; subspinale to right alare

3.1. SAMPLE MATERIAL

The crania used in this research came from three sources: the William M. Bass Donated Skeletal Collection from the University of Tennessee; a collection of skulls located in the Anatomy Museum at the University of Edinburgh; and the Weber and Bookstein Online Material Data from the Virtual Anthropology Textbook (http://extras.springer.com/2011/978-3-211-48647-4). 3D images of crania from the Bass collection and the University of Edinburgh collection were acquired using FastSCAN™ surface laser scans. The Weber and Bookstein data consisted of high resolution CT scans. All specimens were of adult age as determined by the pattern of dental eruption and suture closure and only crania that showed no pathology and no missing landmarks were included in the study. Table 3 provides demographic information, where known, and a list of all the crania utilized. A balanced design with a total of, thirty-one crania of sub-Saharan African or Black American origin and thirty-one crania of European or White American origin was employed.

The William M. Bass Donated Skeletal Collection is housed in the Department of Anthropology at the University of Tennessee, Knoxville in the United States. The collection was established in 1981 and comprises about 1000 individuals with known

demographics including age, sex, and ancestry/race (http://fac.utk.edu/collection.html). The skeletal collection is considered contemporary since most individuals were born after 1940. FastSCAN laser scans of twenty-five skulls were acquired from Professor Caroline Wilkinson and rendered for study from this collection. Twenty-one individuals were of White or Caucasian American origin and four were of Black or African American origin.

Table 3 List and description of cranial samples used

Specimen ID	Collection	Type of Scan	Geographic Origin	Sex (Where Known)	Age
1989 UT02-89D	William M. Bass	Surface Scan	White American	Male	Adult
1992 UT21-92D	William M. Bass	Surface Scan	Black American	Male	Adult
1998 UT05-98D	William M. Bass	Surface Scan	White American	Male	Adult
2001 UT15-01D	William M. Bass	Surface Scan	White American	Female	Adult
2002 UT08-02D	William M. Bass	Surface Scan	White American	Female	Adult
2002 UT35-02D	William M. Bass	Surface Scan	White American	Female	Adult
2003 UT18-03D	William M. Bass	Surface Scan	White American	Female	Adult
2003 UT31-03D	William M. Bass	Surface Scan	White American	Male	Adult
2003 UT49-03D	William M. Bass	Surface Scan	White American	Male	Adult
2003 UT52-03D	William M. Bass	Surface Scan	White American	Male	Adult
2004 UT06-04D	William M. Bass	Surface Scan	White American	Male	Adult
2004 UT18-04D	William M. Bass	Surface Scan	White American	Female	Adult
2004 UT38-04D	William M. Bass	Surface Scan	White American	Male	Adult
2004 UT40-04D	William M. Bass	Surface Scan	Black American	Male	Adult
2004 UT41-04D	William M. Bass	Surface Scan	White American	Male	Adult
2004 UT44-04D	William M. Bass	Surface Scan	White American	Male	Adult
2005 UT10-05D	William M. Bass	Surface Scan	White American	Male	Adult
2005 UT15-05D	William M. Bass	Surface Scan	White American	Male	Adult
2005 UT18-05D	William M. Bass	Surface Scan	Black American	Female	Adult
2005 UT51-05D	William M. Bass	Surface Scan	White American	Male	Adult
2005 UT59-05D	William M. Bass	Surface Scan	White American	Male	Adult
2005 UT79-05D	William M. Bass	Surface Scan	White American	Female	Adult
2006 UT48-06D	William M. Bass	Surface Scan	White American	Male	Adult
2006 UT56-06D	William M. Bass	Surface Scan	White American	Female	Adult
2006 UT75-06D	William M. Bass	Surface Scan	Black American	Male	Adult
XXVI D.9	Edinburgh University	Surface Scan	W. Africa	-	Adult
XXVI D.34	Edinburgh University	Surface Scan	W. Africa	-	Adult
XXVI E.1	Edinburgh University	Surface Scan	S. W. Africa	-	Adult
XXVI E.3	Edinburgh University	Surface Scan	S. Africa	-	Adult
XXVI E.5	Edinburgh University	Surface Scan	Madagascar	-	Adult
XXVI E.7	Edinburgh University	Surface Scan	S. Africa	-	Adult
XXVI E.8	Edinburgh University	Surface Scan	S. Africa	-	Adult
XXVI E.11	Edinburgh University	Surface Scan	S. Africa	-	Adult
XXVI E.12	Edinburgh University	Surface Scan	S. Africa	-	Adult
XXVI E.18	Edinburgh University	Surface Scan	S. Africa	-	Adult
XXVI E.20	Edinburgh University	Surface Scan	S. Africa	-	Adult
XXVI E.22	Edinburgh University	Surface Scan	S. Africa	-	Adult
XXVI E.29	Edinburgh University	Surface Scan	S. Africa	-	Adult
XXVI E.31	Edinburgh University	Surface Scan	S. Africa	-	Adult
XXVI E.36	Edinburgh University	Surface Scan	S. Africa	-	Adult
XXVI F.1	Edinburgh University	Surface Scan	S. Africa	Female	Adult
XXVI F.9	Edinburgh University	Surface Scan	S. Africa	Female	Adult
XXVI F.13	Edinburgh University	Surface Scan	S. Africa	Female	Adult
XXVI F.20	Edinburgh University	Surface Scan	S. Africa	Male	Adult
XXVII 3	Edinburgh University	Surface Scan	Madagascar	-	Adult
XXVII 4	Edinburgh University	Surface Scan	Madagascar	-	Adult
XXVII 5	Edinburgh University	Surface Scan	Madagascar	-	Adult
XXVII 6	Edinburgh University	Surface Scan	Madagascar	-	Adult
VA 001	Weber & Bookstein	CT Scan	Europe	Female	Adult
VA 002	Weber & Bookstein	CT Scan	Europe	Male	Adult

VA 003	Weber & Bookstein	CT Scan	Europe	Male	Adult
VA 004	Weber & Bookstein	CT Scan	Europe	Female	Adult
VA 007	Weber & Bookstein	CT Scan	Europe	Male	Adult
VA 012	Weber & Bookstein	CT Scan	Europe	Male	Adult
VA 014	Weber & Bookstein	CT Scan	S. S. Africa	Female	Adult
VA 019	Weber & Bookstein	CT Scan	S. S. Africa	Female	Adult
VA 021	Weber & Bookstein	CT Scan	Europe	Male	Adult
VA 022	Weber & Bookstein	CT Scan	Europe	Male	Adult
VA 024	Weber & Bookstein	CT Scan	S. S. Africa	Female	Adult
VA 025	Weber & Bookstein	CT Scan	S. S. Africa	Male	Adult
VA 030	Weber & Bookstein	CT Scan	Europe	Female	Adult
VA 050	Weber & Bookstein	CT Scan	Europe	Female	Adult

The crania collection from the Anatomy Museum, housed in the Biomedical Sciences department at the University of Edinburgh, is comprised entirely of individuals of sub-Saharan African ancestry. All crania in this collection are dated to around the nineteenth century. Regional origins are documented for all individuals. A total of twenty-three crania from this collection were scanned and landmarked for this study: fifteen crania came from southern Africa, five from Madagascar, two from western Africa, and one individual from southwestern Africa. Crania were laser scanned by the author using a Polhemus FastSCAN™ portable handheld laser scanner.

FastSCAN™ is a device that acquires three-dimensional surface images of objects by sweeping a handheld laser wand over the object. The wand sweeps a laser light over the object and two cameras on either side of the path swept by the laser record the surface information of the object. The position and orientation of the laser wand is tracked by an electromagnetic field emitted from a transmitter that is securely placed next to the object while it is being scanned. An image of the object instantly appears on a computer screen that is connected to the FastSCAN™ processing unit once the laser beamed line has been swept over the object. When scanning is completed the three-dimensional surface data can be processed and exported in a variety of formats (e.g.,

.ply, which is commonly employed in morphometrics) for use in various software programs. The absolute accuracy of the FastSCAN™ laser scanner is 0.75mm and the practical accuracy is set at 0.13mm. Figure 8 provides a photograph of the device along with a 3D laser surface scanned image of a skull viewed on a laptop.

Figure 8 Polhemus Fastscan Image.

Image of a skull being laser scanned with the Polhemus FastSCAN™ device (left) and the 3D surface scanned image viewed from a laptop. Images were acquired from the Polhemus website (http://polhemus.com/?page=Scanning_Applications_Forensics).

The Weber and Bookstein Online Material Data consist of a collection of CT scans of modern human and great ape crania (Weber and Bookstein, 2011). The human cranial images from this collection used in the present study were all of adult age with documented geographic ancestry. The CT scans of nine European and four sub-Saharan African crania were obtained from the Online Extra Material from the Virtual Anthropology Textbook (Bookstein and Weber, 2011). The program Amira® was used to create three-dimensional models from the .dicom files of the CT scans. This is a software package that is used to visualize three-dimensional image data, primarily volumetric data. The program was used to import and render the CT scanned crania and export the digital images as different file types to be used in morphometric software programs.

Unlike CT scans, laser surface scans only produce image data from the external surface of objects. CT scans include both surface and internal details, creating virtual "solid object" images. To enable landmarking and easier manipulation of the 3D images, the program Freeform® Modeling Plus was used to "fill" and "thicken" the surface scans without losing the morphological integrity of the external surface of the crania.

It should be noted that individuals were not pooled by sex for subsequent analyses since many individuals utilized for the study did not have information relating to this. An individual's ethnic, regional, and national origins were also not considered since some individuals did not have documentation relating to this information. The study can be thought of as 'simulating' a real case scenario in forensic human identification to assess ancestry using the cranium where no prior information is available relating to sex, ethnicity, regional or national origins.

Classification of individuals in terms of ancestry was based on the broad labels, commonly used in forensic anthropology and human identification, relating to the continent/region of origin wherein phenotypic differences were accentuated over time through differing microevolutionary paths. For the purposes of this study, Caucasian American or White crania are referred to under a European ancestry label and African American or Black crania are referred to under a sub-Saharan African ancestry label. The terminology may be debatable or objectionable since some researchers claim that American Whites and Blacks may now be distinct morphologically from modern Europeans and sub-Saharan Africans. It is also not possible to know whether the Caucasian American or African American crania employed from the William M. Bass collection actually have recent European or sub-Saharan roots, as both Whites and

Blacks may have lived in the U.S. for many generations or they may represent more recent immigrants from other ethnically diverse regions other than Europe or Africa. However, one could argue that it is reasonable to assume, that in terms of evolutionary geographic origins, individuals from the William M. Bass skeletal collection, whether White or Black, have at least partial, European or sub-Saharan African ancestry in terms of long term human evolutionary history.

3.2. DATA ACQUISITION

Data used in this research consisted of three-dimensional Cartesian coordinates of landmarks and semilandmarks that were captured on 3D images of crania. Landmarks were placed on three-dimensional virtual images of crania using the program Landmark Version 3.0 (free to download from http://graphics.idav.ucdavis.edu/research/EvoMorph). This program is used mainly for virtual three-dimensional geometric morphometric data acquisition. 3D cranial surfaces were imported into the program and landmark and semilandmark points were placed on the 3D surfaces according to the definition of their locations and in a predetermined sequence so that the relative landmark locations and order would remain constant between all individuals within the study. For landmarks such as ectoconchion and orbitale, whose locations are determined by distance measurements based on the greatest breadth or height of a feature, an on screen desktop measuring application called MB-Ruler was utilized. MB-Ruler is a semi-transparent measuring device that can place a grid on the computer screen as well as measure distances and angles on screen with a triangular ruler (free to download from http://www.markus-bader.de/MB-Ruler/). The on-screen grid was used to aid in the location of landmarks that were based on points of extrema. Once landmark and semilandmarks points were identified and

located on the crania, the raw x, y, z Cartesian coordinate positions of each landmark configuration for each specimen were exported from Landmark and imported into other programs to carry out geometric morpohometric and multivariate statistical analyses.

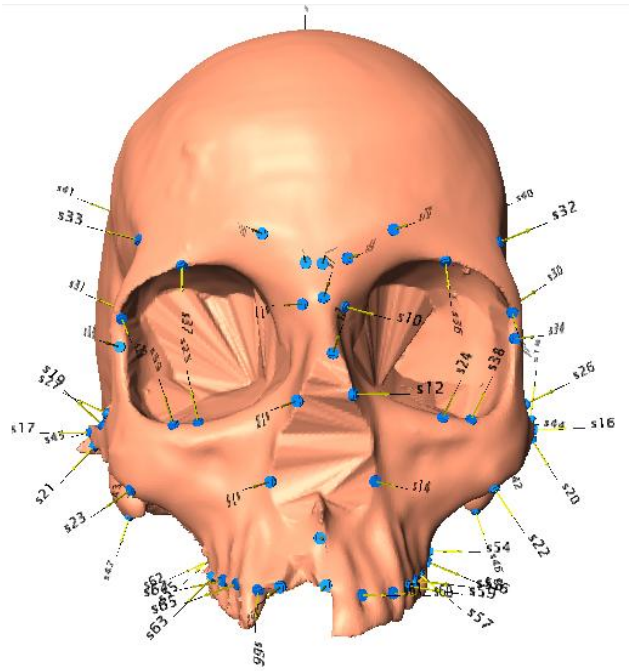

Figure 9 Surface scanned cranium oriented in Landmark program showing landmark points captured

Semilandmarks were digitized along particular features and areas on the skull with the same technique but with variable sample densities from specimen to specimen. The feature being semilandmarked was identifiable across all specimens but the exact location of each point and number could not be defined precisely. For instance, the start point for the orbital curve began at the landmark frontomalare orbitale and semilandmarks were manually placed inferior to this point and traced along the entire orbital rim until they reached frontomalare orbitale from the opposite direction. Individual 2003UT18 had seventy-three semilandmark points arranged along the left orbit whereas individual 2005UT15 had sixty-four semilandmark points positioned on

that feature. Before geometric morphometric methods could be applied to the semilandmark coordinate points, they were resampled so that their density and approximate relative positions would remain consistent among all individuals. The semilandmark coordinates were subjected to a resampling procedure using the MS-DOS *resample.exe* program written by David Reddy and Johann Kim and reprogrammed by Dr. Ryan Raaum (free to download from http://www.nycep.org/nmg/programs.html). The program resampled the semilandmark coordinate data to a specific number of equidistantly-spaced points determined by measuring the total length of the particular curve and dividing it by the desired number of points along the feature.

It should be noted here that if semilandmark points are employed in a geometric morphometric study it is highly recommended that they be algorithmically moved or iteratively 'slid' as an extension of the Generalized Procrustes Analysis process to improve the correspondence between points and among specimens (Bookstein, 1997; Perez et al., 2006). By employing this method, semilandmarks are allowed to 'slide' optimally along areas perpendicular to the surface until their positions are matched, as well as possible, to the corresponding positions of points to a reference configuration (Perez et al., 2006). In this study, semilandmarks, although resampled, were not 'slid' and were treated the same as landmarks. Although sliding is considered more desirable, software available for sliding 3D semilandmarks is currently very limited. Perez and colleagues conducted a study in which they tested two different methods for sliding semilandmarks as well as the effect of not sliding and they found that, in general, differences due to sliding between methods and those semilandmarks not 'slid' had a modest effect on relative shape differences as well as results obtained from discriminant function analyses (Perez et al., 2006).

Dr. Andrea Cardini, a co-supervisor on the project, was able to test ‘sliding’ versus not ‘sliding’ on the semilandmark data from the vault (A. Cardini, personal communication, June 2011). When the semilandmarks on the vault were Procrustes superimposed with object symmetry, asymmetries were removed and the third dimension became obsolete for this dataset. Removing the third dimension made the semilandmarks 2D. The semilandmarks from this dataset were ‘slid’ using TPSRelw, a geometric morphometric program which provides a procedure to slide 2D semilandmark data. Shapes distances between ‘sliding’ semilandmarks and those not ‘slid’ were shown to be negligible with virtually no change (results not shown) (A. Cardini, personal communication, June 2011). Results obtained should, hence, be reliable even if semilandmarks were not ‘slid’ for other datasets since distances did not change appreciably by ‘sliding’ for the semilandmarks along the vault. The main aim of the project is classification accuracy of landmark and semilandmark configurations using discriminant function analyses. Results of statistical analyses should generally be small when semilandmarks are not ‘slid’ since discriminant analyses rely on distances of individuals from the mean of the groups

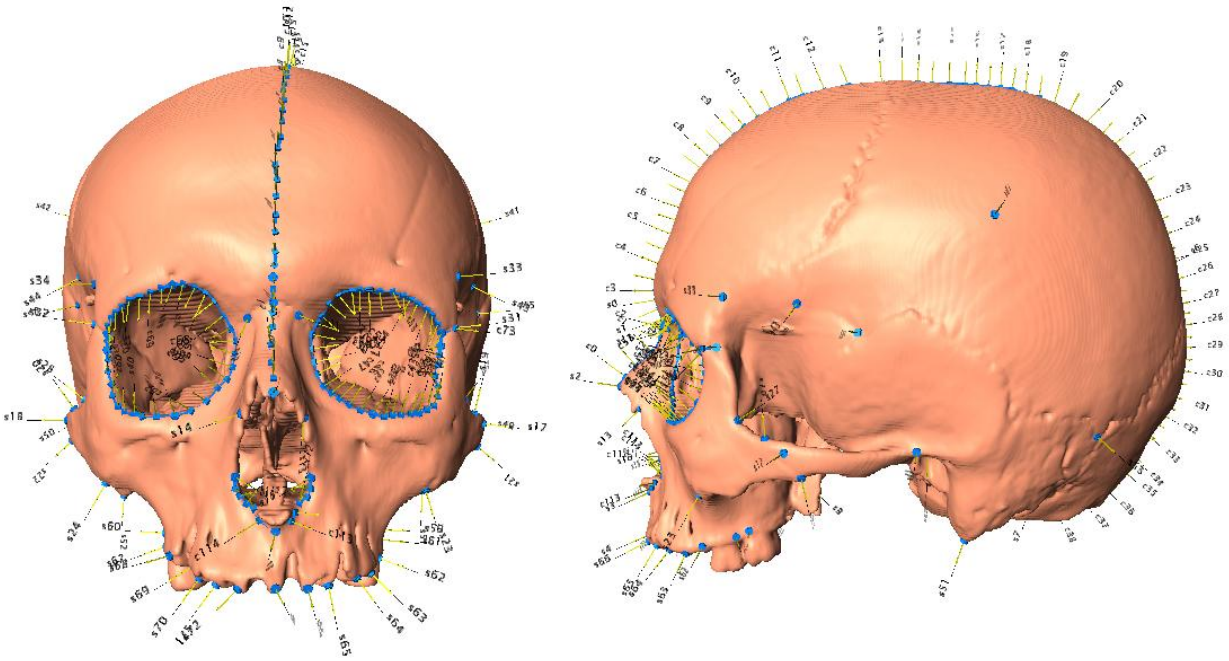

Figure 10 CT scanned cranium oriented in Landmark showing landmark and semilandmark points captured in frontal and profile positions

3.3. DATA ANALYSIS

3.3.1. SHAPE ANALYSES

Procrustes-based geometric morphometric methods were used to analyze and compare the coordinate configurations of landmarks and semilandmarks collected from the 3D cranial surfaces. The relative positions of the landmarks and the predetermined order in which they were selected and located ensured that there would be a one-to-one correspondence among landmarks across all specimens used in the study (Klingenberg, 2011). The program MorphoJ Version 1.00j was used to perform all geometric morphometric analyses on all the landmark and semilandmark coordinate data exported from the Landmark program (free to download from http://www.flywings.org.uk/MorphoJ_page.htm).

Analyses were carried out using the entire dataset (consisting of all the cranial landmarks and semilandmarks), as well as using only subsets of the data. The different subsets used included (1) conventional landmarks recorded over the entire cranium, (2) only semilandmarks recorded along the vault in the midplane, (3) only semilandmarks recorded along the orbits, (4) only semilandmarks recorded along the lower nasal aperture, and, as mentioned, (5) the total set of points. Shape information from the landmark and semilandmark coordinates was obtained in MorphoJ by a Generalized Procrustes Analysis (GPA). GPA was used to register the coordinates and fit all individuals into a common coordinate system from which differences in size, translation and position are 'standardized'. An average mean shape was produced and the landmark points captured from each specimen were translated to a common origin, rescaled to unit centroid size, and rotated relative to one another to minimize the sum of squared distances between corresponding coordinates for all specimens (Rohlf and Slice, 1990; Harvati, 2009; Weber and Bookstein, 2011). The centroid size, measured in centimetres, is a measure of dispersion around the centroid of the points in a landmark/semilandmark configuration. Once GPA was performed and the points were brought into a common coordinate system the differences between the values of the coordinates of the landmarks measured the differences in shape between the specimens (Slice, 2007). Small left-right asymmetries were removed following Klingenberg et al. (2002). GPA shape coordinates were subjected to statistical procedures to reduce the number of variables within the dataset, to distinguish between shape similarities and differences among and between groups, and to classify groups using discriminant function analyses.

3.3.2. STATISTICAL COMPARISONS

A Principal Components Analysis (PCA), using the variance-covariance matrix of the GPA shape coordinates, was performed in MorphoJ as well as in the statistical software PAST.exe to obtain new variables to summarize and explore visually the multivariate pattern of shape variation in the total sample. PCA comes from applied linear algebra and is an investigative multivariate statistical technique that provides a means to ordinate complex data and summarizes the variation occurring within it by creating new variables from the original data by re-expressing variables in a different way (Shlens, 2005; Cardini, 2012). The dimension of the complicated data set is reduced by transforming the data to a new set of variables by extracting eigenvectors from the covariance matrix (Zelditch et al., 2004). The projections of the datapoints on the eigenvectors are called the principal components (PCs), which are linear combinations of the shape variables and are independent of one another (Zelditch et al., 2004). The PC's are uncorrelated (i.e. independent of each other) and are sorted so that the first principal component represents the largest amount of shape variation within the sample, and the second, third, fourth components, etc. represent progressively decreasing amounts of shape variation until all significant variation is accounted for. In the PCA scatterplots, distances between the specimens are proportional to differences in shape along the visualized axes. Scatterplots of the PCA were produced for the different datasets and the specimens were arranged in a shape space so that similarities and differences between individuals could be visualized.

Dimensionality reduction removes potentially redundant variables and it may be particularly desirable when utilizing certain statistical techniques such as discriminant function analysis to assign individuals into groups. Discriminant function analyses do

not perform well with small sample sizes and many predictor variables, which can be common to geometric morphometric studies, especially when semilandmarks are analysed (Kovarovic et al., 2011).

To test whether the landmark and semilandmark coordinates captured from the different datasets were able to classify individuals according to their designated ancestry group, a series of discriminant function analyses were carried out. One used only centroid size, a second used the principal component shape scores, a third used the principal component form scores, and a fourth used the principal component form scores but with the first principal component removed to control for size-common allometric differences within the samples (i.e. 'size-corrected' form). Form scores were obtained by performing a PCA on the matrix of shape coordinates to which the natural log of centroid size was appended. This generates a size-and-shape space (or Procrustes form space) where variation in both types of form differences can be simultaneously analysed (Mitteroecker and Gunz, 2009).

The discriminant function analyses were performed on the selected set of predictors (size, shape, form, and 'size-corrected' form) using leave-one-out cross validation in SPSS for Windows version 13 as well as in PAST. In total, twenty discriminant analyses using leave-one-out cross validation were performed using the principal component scores that accounted for at least 95% of the variation within the dataset. Discriminant analyses based on form and "size-corrected" form included one additional principal component to account for the extra dimension of size that was used.

The leave-one-out cross validation method assessed the predictive performance of the discriminant function by testing the accuracy of each predicted classification. By using leave-one-out cross validation in the discriminant function analyses each

individual was iteratively removed from the whole sample and treated as an unknown for the discrimination procedure (Polly and Head, 2004). The identity of each individual was estimated by using the rest of the sample specimens whose identities were known (Polly and Head, 2004). The leave-one-out cross validation procedure provides an unbiased estimate of the probabilities of correct classification and diminishes issues of overfitting common to discriminant analyses done without cross-validation (Buck and Vidarsdottir, 2004; Polly and Head, 2004; Kovarovic et al., 2011). The landmarks and semilandmarks captured were assessed by their predictive power in classifying ancestry correctly depending on the variable/predictor used.

Posterior probabilities of group membership were also calculated in the discriminant function analysis. Posterior probabilities are the measured likelihood that a given individual belongs to either group. Posterior probabilities, however, are calculated with the assumption that all individuals within the sample belong to one of the groups in the analysis (Pietrusewsky, 2000). With two balanced groups, individuals correctly classified have a posterior probability greater than 0.5 and those individuals incorrectly classified will have a posterior probability less than 0.5. Typicality probabilities, which assess the likelihood that any given individual actually belongs to any of the groups assigned, were also calculated for two datasets using a single predictor. The dataset that analyzed the shape of cranial landmarks and the size from the lower nasal aperture were used to evaluate the likelihood that any given individual belonged to any or none of the groups represented. Typicality probabilities should have been calculated for all datasets and variables, however, due to time constraints this was not possible. The typicality probabilities computed for the two datasets provide an idea, at least, of how typical individuals are to the mean of the group in the discriminant function analyses. Typicality probabilities are calculated by determining the average variability of all

groups in an analysis and measures range from 0 to 1 (Pietrusewsky, 2000). Generally, a low typicality (for this study, set at less than 0.01, the same threshold conventionally used for 'high significance') suggests that an individual may not belong to any of the groups within the analysis and may possibly belong to another group entirely. The lower the typicality value the further that individual is from the mean of the group in the discriminant analysis function. Since posterior probabilities are relative probabilities, and typicality probabilities are absolute probabilities, an individual could have a very high posterior probability but an extremely low typicality probability. In this case, the individual is closer to a given group but so distant from the average that the individual could actually represent an outlier (e.g., a third group in terms of ancestry).

Kovarovic and colleagues discuss the random chance baseline and the importance of calculating a 'chance-corrected' classification statistic in discriminant function analyses (2011). This calculation provides information about the proportion of individuals correctly classified into respective groups by chance alone. When the sample size is completely balanced between groups this value is $1/G$, where G represents the number of groups within the analysis. Within this study ancestry groups are balanced (there are 31 sub-Saharan Africans and 31 Europeans) so the expected proportion of cases correctly classified by pure chance alone expressed as a percentage is 50%. The TAU statistic takes this prior probability and removes the probabilities due to chance from the correct classification rates obtained from the discriminant analysis. (Kovarovic et al., 2011). The formula to calculate this statistic can be found in Kovarovic et al. (2011). Values of zero signify that the correct classification rate is no better than if individuals were grouped randomly by chance and values of one indicate that there was perfect discrimination (Rodríguez-Mendoza et al., 2011). The TAU values were calculated for each dataset and each predictor variable and were expressed

as a percentage and compared to the correct classification rates obtained from the cross-validated discriminant function results.

An ancestry bias index was also calculated following the methods proposed by Franklin and colleagues (2012). The ancestry bias is the difference in correct assignment of sub-Saharan African and European individuals from the total sample (Franklin et al., 2012). In the present study, a positive value signifies that more sub-Saharan African individuals were misclassified relative to European individuals and a negative value indicates the opposite.

3.4. Error analysis studies

3.4.1. Accuracy of landmarks on 3D images vs. those directly digitized from bone

A study was performed to explore and compare the amount of variation between cranial landmarks acquired from 3D images and those acquired directly by a digitizing device in relation to inter-individual (i.e., observed sample) variance. The crania utilized for this purpose came from the Helmer collection and the Centre for Anatomy and Human Identification teaching collection both housed at the University of Dundee. 3D images of ten crania were landmarked virtually using Landmark software and the same ten crania were directly digitized with the same set of landmarks using a Polhemus Patriot™ Digitizer.

The Patriot™ digitizer is a device that can capture the real-time 3D Cartesian coordinate points of an object. The crania directly digitized with the Patriot™ were mounted and fixed in position to enable landmarking. The degree of accuracy of the Patriot™ itself was tested by taking a series of points along a ruler at one centimetre intervals. The Patriot's components generate an electromagnetic field that can detect the

position and orientation of the sensor on the tip of a stylus pen when it is placed within the field. The x, y, z points collected from the position of the stylus tip are recorded into an associated software program installed onto a PC.

Thirty-three midline and bilateral landmark points, representing Type I, Type II, and Type III landmarks, were directly digitized as well as captured from the 3D images and were subjected to GPA in MorphoJ. Table 4 provides a list of the landmarks used for this particular error study. Definitions for landmark placement can be found in Table 1.

Table 4 Cranial landmarks used for digitization error study.

Order placed	Landmark name
1	Glabella
2	Nasion
3	Rhinion
4	Subspinale
5	Prosthion
6	Bregma
7	Lambda
8	Opisthion
9	Basion
10 and 11	Infranasion
12 and 13	Naso-maxillary
14 and 15	Alare
16 and 17	Inferior zygotemporale
18 and 19	Jugale
20 and 21	Frontomalare orbitale
22 and 23	Frontomalare temporale
24 and 25	Frontotemporale
26 and 27	Auriculare
28 and 29	Mastoidale
30 and 31	Lateral foramen magnum
32 and 33	Ectomalare

The Procrustes shape coordinates were analyzed using a Procrustes analysis of variance (ANOVA) to quantify digitization and measurement error between instruments (Klingenberg and McIntyre, 1998). The Procrustes ANOVA test simultaneously determined the variation in the sample according to various shape effects and size differences between individuals as well as both the differences between landmarks acquired using either 'instrument' and the overall effect of measurement error (i.e. repositioning of the specimen). The Procrustes ANOVA in MorphoJ is tailored to analyze patterns of asymmetric/symmetric shape variation in data and therefore, besides testing measurement error, provides tests for the significance of the asymmetric components of shape variation. This involves four steps: (1) a calculation of the amount of total shape variation between individuals within the dataset; (2) a calculation of the amount of directional asymmetry among individuals, which provides information about whether one side of the specimen that was landmarked is systematically larger than the other side; (3) a calculation of the amount of fluctuating asymmetry, which represents small deviations from absolute bilateral symmetry (i.e.; variation between left-right landmarks) (4) a calculation of the amount of variation that is left over (the residual variation) within the dataset, which provides a value for measurement error within the dataset of replicated digitizations of landmarks (Klingenberg and McIntyre, 1998; Singh et al., 2012).

3.4.2. Digitization error in 3D images

A landmark error study was also performed to see if the cranial landmarks placed on the 3D images used in the main analysis were repeatable throughout the study. Five crania from the original study sample were selected and re-landmarked on three separate occasions with the complete set of sixty-eight cranial landmarks. Repeat

landmarking sessions were carried out with at least a one-week interval between them to prevent measurement bias through remembered specimens and remembered landmark positions. Landmark precision was not tested for each individual landmark coordinate but on the set of landmarks as a whole. The aim of this intra-observer landmark precision study was similar to the Polhemus versus 3D error test but was conducted on the 3D images of crania with replicated digitizations of landmarks only. In total, four replicates of five specimens were utilized since landmarks from the original study were also included in this test of measurement error. Landmark precision was assessed between the re-landmarked specimens in relation to the natural variation between different individuals within the sample.

Measurement variability was visualized by employing a GPA and PCA on the sets of the standard cranial landmark coordinates with the inclusion of the re-landmarked specimens following the methods proposed by O'Higgins and Jones (1998). Procrustes ANOVA was conducted on the shape coordinates of the error dataset to calculate the variation due to measurement error in comparison to the variation between individuals within the sample. PCA was performed to provide a visualization of how the replicated landmark configurations plotted in relation to one another and to the entire dataset of all individuals from the main analysis. Because of the length of the digitization procedure for semilandmarks, these were not included in the replicas and therefore in the error analysis. However, with consideration (see Results) to the large amount of inter-individual variation in the sample, the analysis based on the landmarks should be valid and its results should apply to semilandmark data as well.

4. RESULTS

Reported first, are the error tests: (1) a comparison of digitizing error between digitizing 'instruments' (acquisition of 3D image landmarks vs. Polhemus Patriot landmarks), and, (2) landmark measurement error and repeatability on 3D images as a whole. Following this, visualizations of some of the PCAs are explored by providing scatterplots relating to some of the datasets along the first few PCs to view shape and form interactions among and between individuals and groups. Shape changes were also visualized in the Landmark software for the cranial landmark set along the first PC shape change vector for extreme European and extreme sub-Saharan African shapes as well as visualized as a mean group shape from either ancestry group. These were exported as coordinate files exported from MorphoJ. Classification results of the discriminant function analyses on the various landmark and semilandmark datasets are also presented with the inclusion of the ancestry bias index, posterior probabilities and typicality probabilities (the latter only for the dataset of cranial landmarks using the variable of shape and semilandmarks along the edge of the lower nasal aperture using the variable of size).

4.1. 3D IMAGE LANDMARKS VS. POLHEMUS PATRIOT LANDMARKS

Landmark placement variation between 'instruments', or digitizing error, were quantified relative to the total shape and size variation among individuals within the dataset using Procrustes ANOVA. Tables 5 and 6 present the results of this test. The Procrustes ANOVA summarizes the overall effects of digitization error versus the total shape and size variation of individuals within the dataset. The degree of digitization error for the Procrustes ANOVA is calculated in relation to the degree of individual shape and size variation. Results show that the level of error between digitization

methods and, hence, measurements with either device, relative to inter-individual shape and size interactions, is not statistically significant. Individual shape and size interactions were statistically significant (probability value, or p=<.0001) compared to measurement error. This indicates that digitization of cranial shapes from 3D images is comparable to landmarks digitized directly from crania since the amount of error between replicates is significantly less than the amount of shape and size effects explained between individuals

Table 5 Procrustes ANOVA results for 'instrument' and measurement error using centroid size

Effect	SS	MS	df	F statistic	p-value
Individual	7.979642	0.886627	9	24.07	<.0001
Residual	0.368341	0.036834	10	-	-

SS, MS, and df refer respectively to sum of squares, mean sum of squares (i.e., SS divided by df) degrees of freedom. Residual effect refers to 'measurement error' between replicates in the dataset.

Table 6 Procrustes ANOVA results for 'instrument' and measurement error partitioning symmetric and asymmetric components of shape.

Effect	SS	MS	df	F statistic	p-value
Individual	0.051903	0.000115	450	15.47	<.0001
Side	0.000521	0.000012	42	1.67	0.0076
Individual * Side	0.002818	0.000007	378	0.56	1.0000
Residual	0.012321	0.000013	920	-	-

4.2. 3D IMAGE ERROR

Landmark precision in the study was calculated by re-landmarking five 3D cranial images from the main analyses on three different occasions. All sixty-eight landmarks from the standard landmark set from the main analysis were used. The magnitude of individual shape and size variation was compared to the variation between landmarks placed on replicated individuals. Results of the intraobserver error measurement study for landmark digitization are presented below in Tables 7 and 8. Inter-individual shape variation was greater than fluctuating asymmetry which was

greater than the replicated configuration of landmarks (i.e. residual or measurement error). Therefore inter-individual variation is also greater than the replicas, indicating that measurement error was not significant source of variation within the sample. Results of the Procrustes ANOVA test indicate that individual variation from the landmark configurations was shown to be significantly larger for all shape and size interactions (p=<0.0001) in comparison to replicated landmark configurations as a whole. This indicates that landmark precision was accurate and repeatable within the bounds of inter-individual variation and that measurement error was negligible overall and was not a significant source of variation within the sample.

Table 7 Procrustes ANOVA results for 3D image error using centroid size

Effect	SS	MS	Df	F statistic	p-value
Individual	3185.066	796.2664	4	2720.77	<.0001
Residual	4.389929	0.292662	15	-	-

Table 8 Procrustes ANOVA results for landmark precision on 3D images.

Effect	SS	MS	df	F statistic	p-value
Individual	0.081378	0.000198	412	9.94	<.0001
Side	0.003964	0.000042	94	2.12	<.0001
Individual * Side	0.007469	0.000020	376	21.67	<.0001
Residual	0.002708	0.000001	2955	-	-

A GPA and PCA were also carried out to visualize how the replicated landmark configurations plotted with the data gathered from the sixty-two study individuals included in the main analysis. Figure 11 shows the first two principal components from this analysis plotted against one another. PC1 against PC2 shows that those individuals with replicated landmark configurations cluster relatively close to one another in comparison to all the other individuals within the sample (O'Higgins and Jones, 1998).

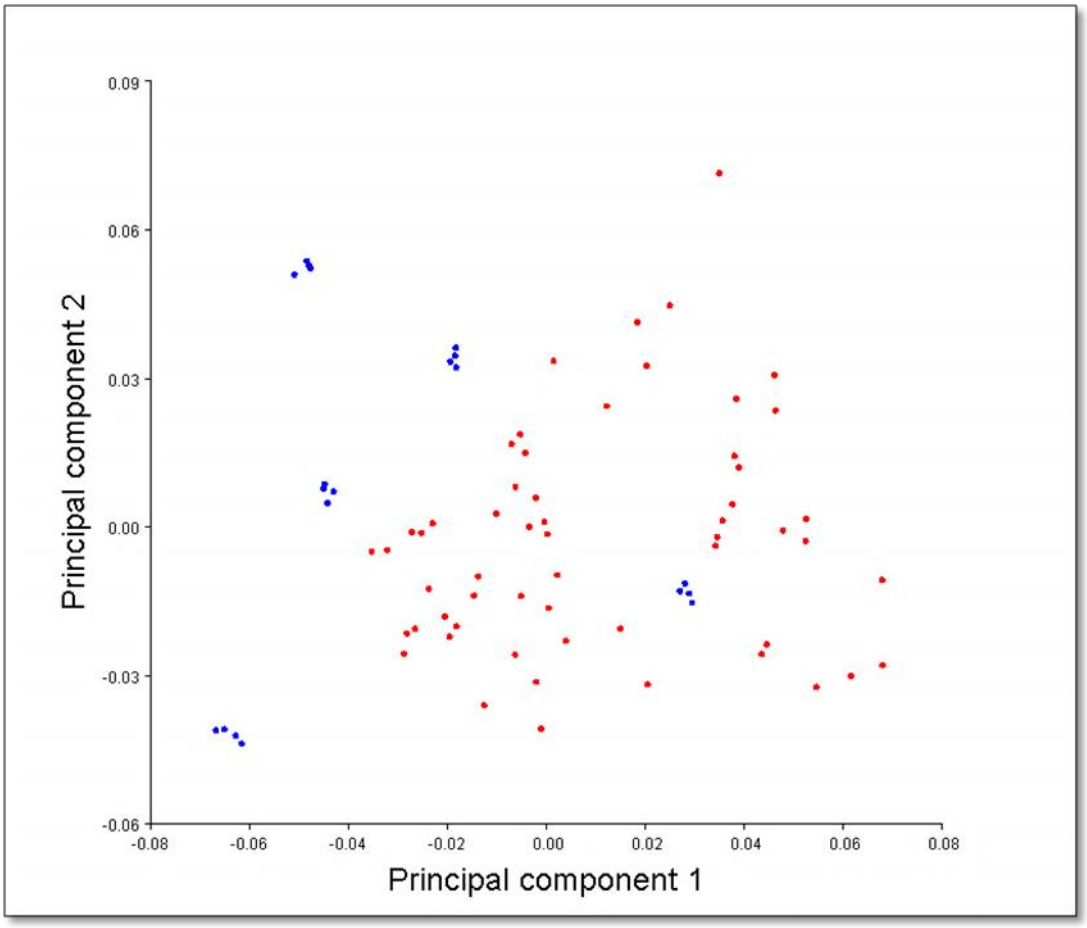

Figure 11 PCA of 3D image digitization error.

The three replicates of each of five individuals together with the data collected from the sixty-two individuals from the full analysis are plotted along PC 1 and PC 2 from the PCA. The replicated landmark configurations/individuals are shown as blue dots and the remaining unreplicated landmark configurations/individuals are shown as red dots. The replicates cluster relatively closely together in comparison to the variation present between individuals within the sample.

4.3. PRINCIPAL COMPONENT ANALYSES AND VISUALIZATION OF DIFFERENCES BETWEEN GROUPS

Raw coordinates from the various datasets utilized were imported into MorphoJ and subjected to Procrustes superimposition as well as PCA in MorphoJ and PAST to obtain variables relating to size, shape, and form from the landmark and semi landmark points.

Size was not a major contributor in terms of overall differences between individuals of sub-Sahara African and European Ancestry. Mean sizes of the groups using the standard landmarks, orbit, and vault dataset as well as the complete

configuration of landmarks and semilandmarks set had almost equal means and the distribution of size between the two groups almost completely overlapped. Size differences, however, are evident for the lower nasal aperture between groups but not for the cranial landmarks. Although there is some overlap between the two groups, mean centroid size from the lower nasal aperture is different between sub-Saharan Africans and Europeans. Figure 12 shows an example of the size variation grouped by ancestry for the cranial landmarks and lower nasal aperture datasets.

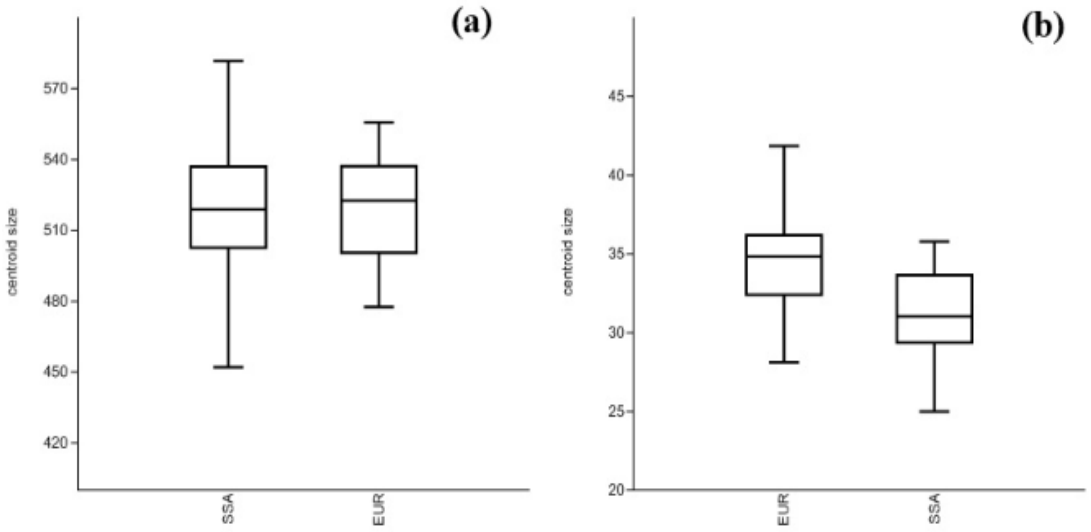

Figure 12 Box plots depicting size variation between ancestry groups from (a) the standard cranial landmark dataset and (b) semilandmarks arranged along the lower nasal aperture. EUR = European individuals; SSA = sub-Saharan African individuals.

Shape variation scatterplots of the first two PCs of the PCA for individuals from the standard cranial landmark set, the semilandmarks on the lower nasal aperture, semilandmarks along the vault, and semilandmarks along the orbits are displayed in Figures 13, 14, 15, and 16, respectively. The first two PCs for these three datasets accounted for the following percentages of variation between individuals: standard cranial landmark set, 37.6%; lower nasal aperture, 34.9%; orbits, 65.4%; and vault, 55.0%. The PCA on the standard landmark dataset for shape generated sixty-one PCs that explained 100% of the variation within the dataset. The first twenty-nine PCs represented at least 95% of the shape variation between individuals and were utilized as the shape variables to predict ancestry for the discriminant function analysis. In this and the other datasets, selecting the first PCs whose cumulative variance explained 95% of total variance allowed the dimensionality to be reduced without changing the pattern of pairwise individual distances in the shape space, as suggested by matrix correlation between Euclidean distances in the subspace of the first PCs and Procrustes distances in the total shape space as large as or larger than 0.997. For the lower nasal aperture eleven PCs were generated and the first four were used for classification analyses. Figure 13 and Figure 14 show that individuals from the standard landmark configuration and the configuration of semilandmarks from the lower nasal aperture tended to plot, in terms of ancestry groupings, along the first and second PCs with some overlap between the groups. PCs 3 and 4, however, show no distinct ancestry groupings with almost complete overlap between groups.

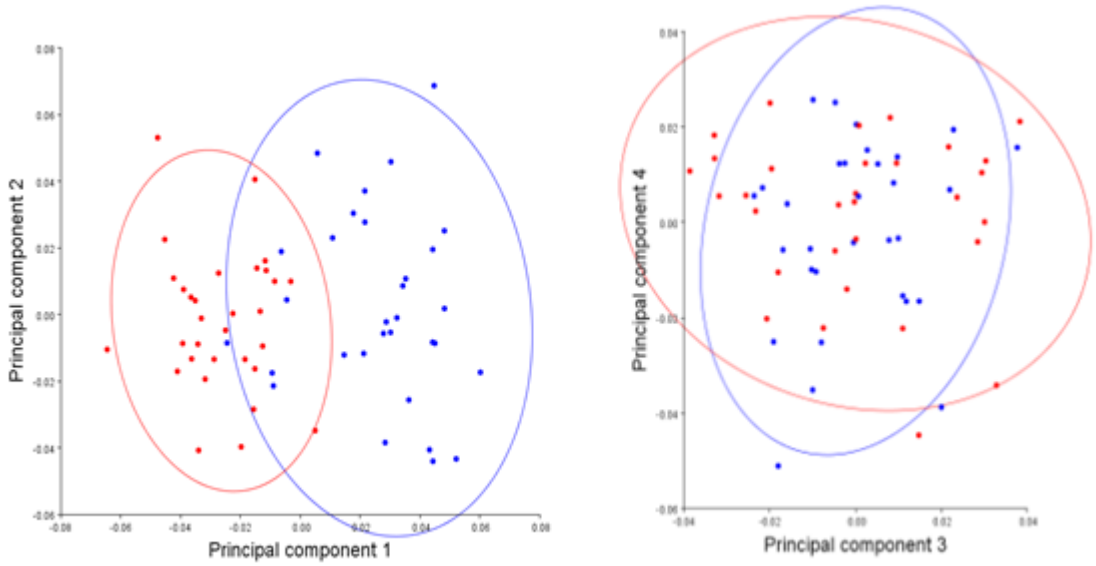

Figure 13 Shape variation between groups from the standard landmark dataset with 95% confidence ellipses. Scatterplots of PC1 vs. PC2, which explain 37.6% of variance, and PC3 vs. PC4, which explain 15.6% of the variance. Blue dots represent sub-Saharan African individuals and red dots Europeans.

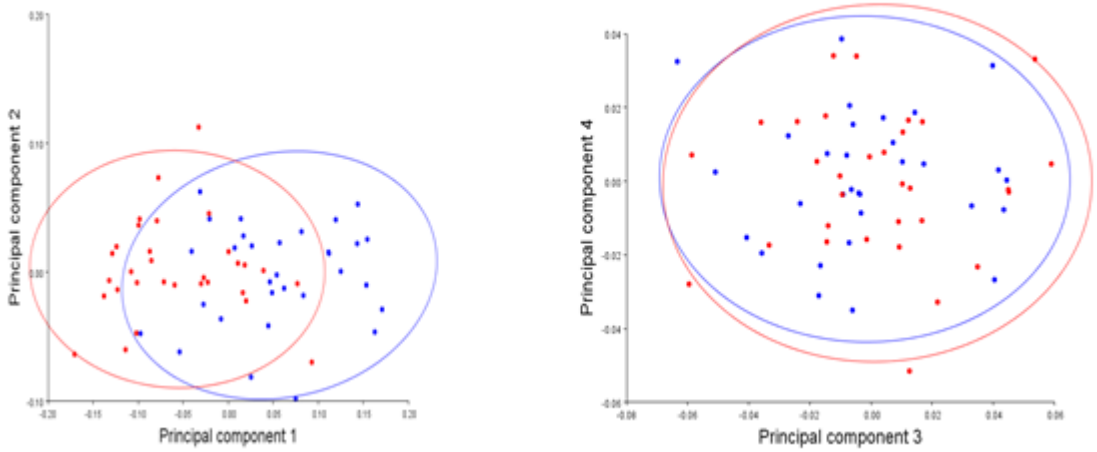

Figure 14 Shape variation between groups from the lower nasal aperture dataset with 95% confidence ellipses. Scatterplots of PC1 vs. PC2, which explain 34.9% of variance and PC3 vs. PC4, which explain 4.2% of variance. Blue dots represent sub-Saharan African

For the PCA of semilandmarks along the vault and those along the orbital rims, forty-seven PCs and forty-nine PCs, respectively, were produced that explained 100%

of the variation among individuals within each dataset. The first nine PCs were used as shape variables for the discriminant analysis on the vault and the first eight PCs were used for the classification analysis on the orbits.

Figure 16 and Figure 16 display the scatterplots of the first two PCs of shape from the vault and orbit. Much overlap, which is especially visible within the orbital dataset, occurs between individuals from either group.

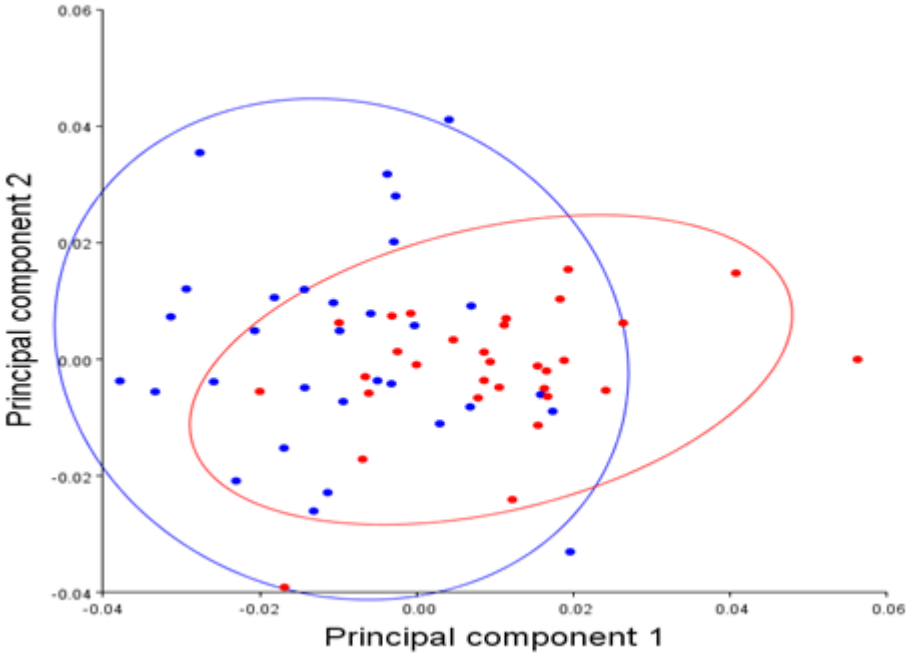

Figure 15 Shape variation between groups from the vault dataset with 95% confidence ellipses. Scatterplot of PC1 vs. PC2 which explain 55.0% of variance. Blue dots represent sub-Saharan African individuals and red dot Europeans.

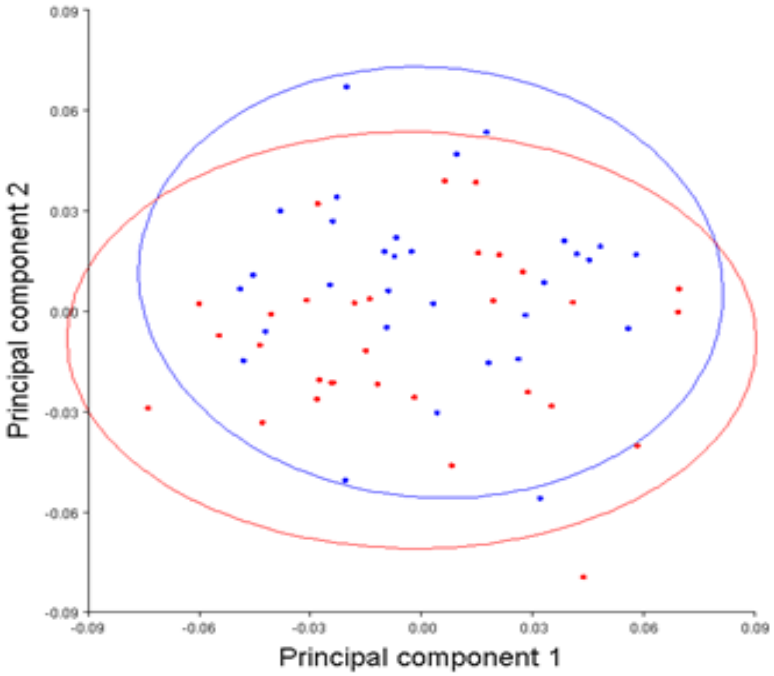

Figure 16 Shape variation between groups from the orbit dataset with 95% confidence ellipses. Scatterplot of PC1 vs. PC2 which explain 65.4% of variance. Blue dots represent sub-Saharan African individuals and red dots Europeans.

Scatterplots of the first two PCs of form and 'size-corrected' form are depicted for the standard cranial landmark set and the lower nasal aperture dataset in Figure 17 and Figure 18, respectively. Form and 'size-corrected' form plots for the standard landmark configurations show a separation of individuals in terms of ancestry group with some overlap between groups. The first two PCs of form and 'size-corrected' form accounted for 55.3% and 21.5% of the variation within the cranial landmark dataset.

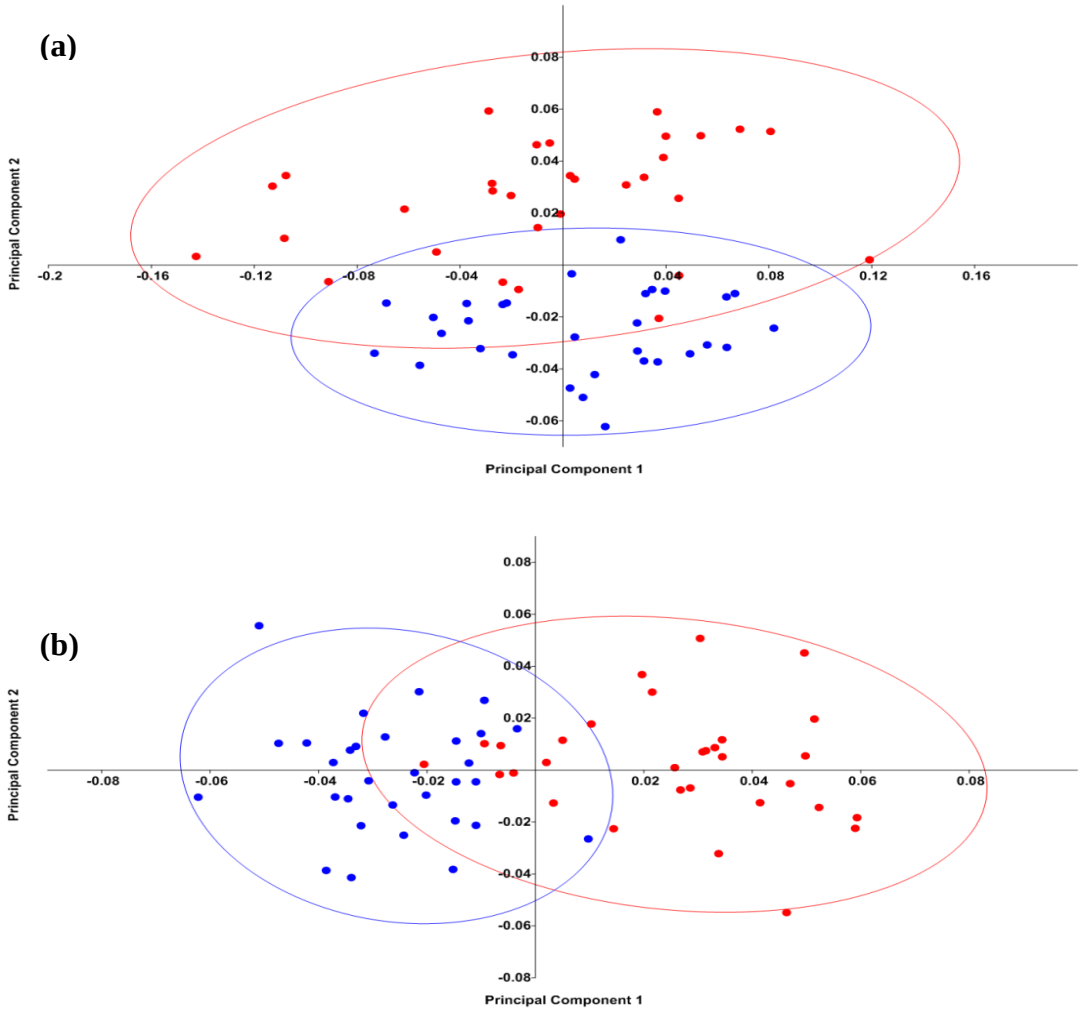

Figure 17 Form (a) and 'size-corrected' form (b) variation between groups from the cranial landmark dataset with 95% confidence ellipses. Scatterplots of PC1 vs. PC2 explain 55.3% of form and 21.5% of 'size-corrected' form variance. Blue dots represent sub-Saharan African individuals and red dots Europeans.

The first two principal components from the lower nasal aperture form and 'size-corrected' form space accounted for 86.4% and 39.2% of the variation within datasets. As a whole, the lower nasal aperture form space separated groups in terms of ancestry with some overlap between groups. The 'size-corrected' form space, however, displayed a greater deal of overlap between groups with no obvious separation between ancestry groups.

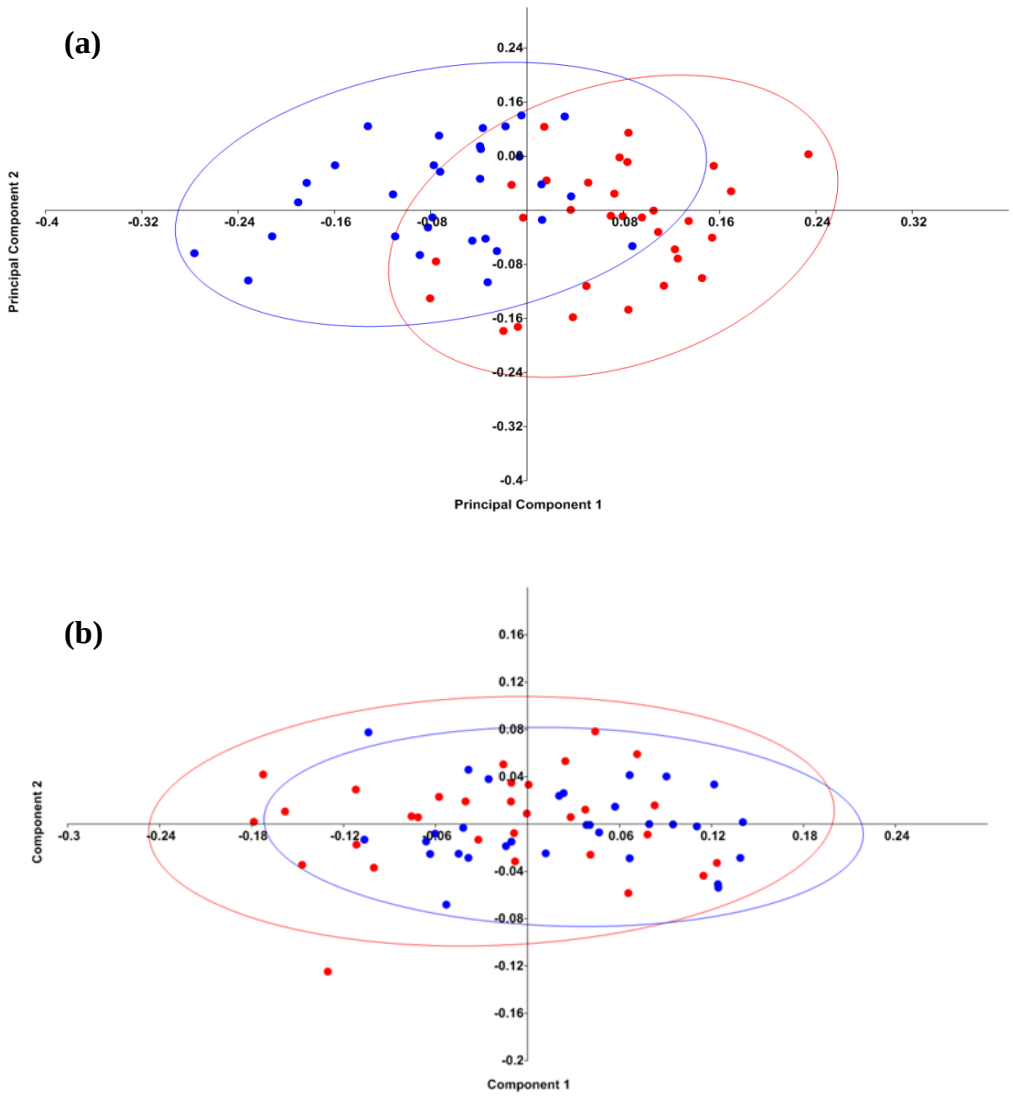

Figure 18 Form (a) and 'size-corrected' form (b) variation between groups from the lower nasal aperture dataset with 95% confidence ellipses. Scatterplots of PC1 vs. PC2 explain 86.4% of form and 39.2% of 'size-corrected' form variance. Blue dots represent sub-Saharan African individuals and red dots Europeans.

Currently there are few programs available to visualize shape changes from 3D landmark data together with the 3D surfaces on which they were digitized. The Landmark program, however, offers a means by which to visualize shape changes from landmark coordinates by rendering surfaces along principal component vectors obtained from MorphoJ. The rendering and the shape changes on the regions of the surface where there are no landmarks is achieved by using the thin-plate spline interpolation (Adams et al., 2002). Figure 19 provides a visualization of an unaltered cranial 3D image of a European individual from the main sample morphed along the first PC of shape variation generated from MorphoJ from the standard landmark configuration with the dataset subdivided into a PCA on the extreme axis of the European group and a PCA on the extreme axis of sub-Saharan indviduals alone. Although interpretation is subjective when analyzing a 3D deformation of this sort, it is helpful and effective to visualize the shape changes occuring since it can provide an intuitive aid for the interpretation of shape variation along a vector.

The image of the crania morphed along the first PC in the direction of the sub-Saharan Africans displays an overall shortening of the crania specifically in the maxillary region when compared to the unaltered 3D cranial image. The frontal area is also slightly shorter and more rounded. Orbital shape is less round and more angular and orbits slope downwards. The nasal opening appears rounder and wider and the nasal bridge lower and flatter with little overhang of the nasal bones. The mastoid process is short and rounded.

The crania morphed along the first PC in the direction of the Europeans displays a longer and narrower cranium and the maxillary region appears longer than in the unaltered image. The frontal bone is high and and more rounded. Orbits are more angular and square-like. The nasal aperture is long and narrow and the nasal bridge is narrow and sharp with overhanging nasal bones. The mastoid process is large, pointed, and long

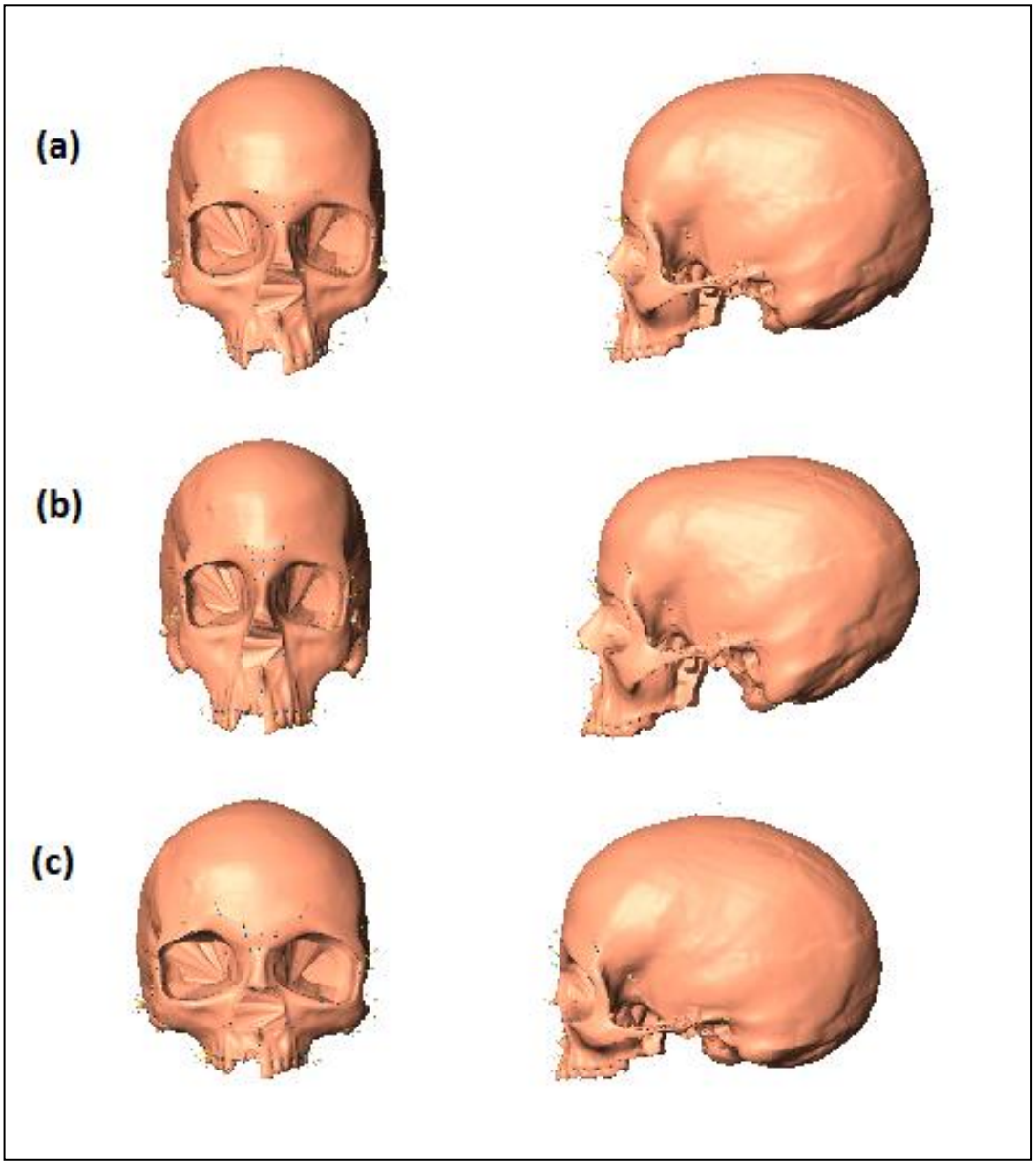

Figure 19 Crania morphed along the first principal component of a PCA on shape; (a) is an unaltered cranial image, (b) is the same cranial image but morphed along the first principal component at the extreme axis of European individuals, and (c) is the same cranial image but morphed along the first principal component at the extreme axis of sub-Saharan African individuals

A visualization of the mean shape of sub-Saharan and European individuals using the cranial landmark dataset is displayed in Figure 20 and magnified three times to easily perceive differences in cranial shape. Mean shape of European crania is similar to the morphed shape using the first PC for that group although shape of the cranium, in general, does not display much difference from the unaltered image; orbits are angular and slope downwards laterally; nasal aperture changes little from the unaltered image but in general is long and narrow; nasal bones overhang the nasal aperture; and mastoid process is long and pointed. The mean shape of the sub-Saharan African crania is also similar to the morphed shape using the first PC for that group however noticeable facial projection/prognathism in the alveolar region and a wider palate is evident. The frontal area is also straighter. From the profile, the cranium is longer laterally. Orbits are large and slope downward laterally; nasal aperture is shorter and wider; and mastoids are rounder and less pointed.

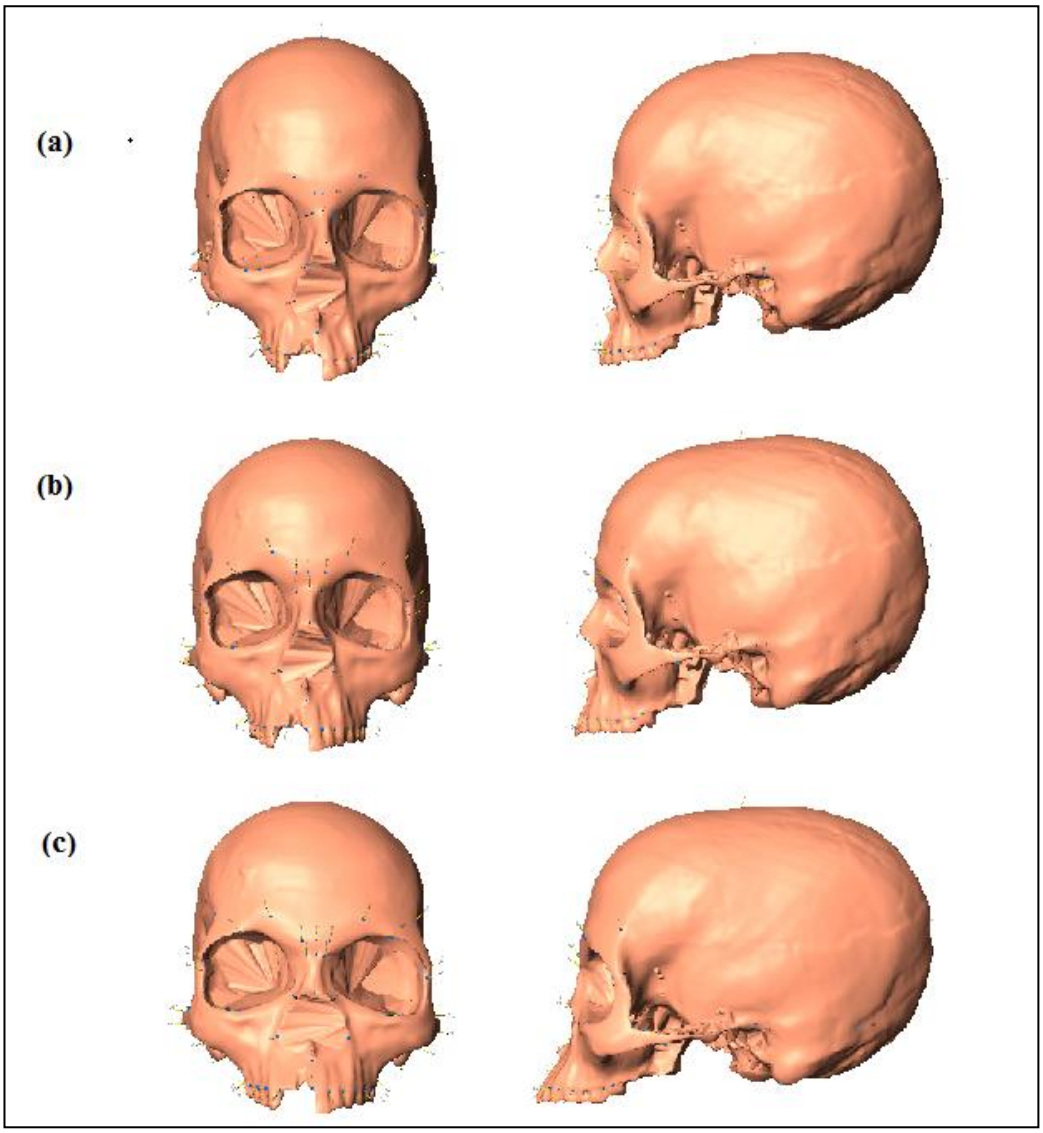

Figure 20 Crania morphed into the mean shape of ancestry groups; (a) is an unaltered cranial image, (b) is the same cranial image but morphed into the mean shape of European individuals, and (c) is the same cranial image but morphed into the mean shape of sub-Saharan African individual

4.4. DISCRIMINANT FUNCTION ANALYSES; CLASSIFICATION ACCURACY AND CONFIDENCE

Results of the discriminant analyses are presented for five datasets: standard cranial landmarks, semilandmarks along the cranial vault along the midplane, semilandmarks along the orbital rims, semilandmarks on the outline of the ridge of the lower nasal aperture, and the combined total of all landmarks and semilandmarks. Variables used for the various discriminant function analyses were centroid size, the PC scores of shape, and the PC scores of form. Analyses of form used the same number of PCs as the shape-only analysis and one additional PC was utilized to account for the variable of size that was included. 'Size-corrected' form used the same number of PCs as for the form analysis but with the first PC omitted to control for the effect of allometry among individuals since the first PC in a form analysis is a common allometric variable. All discriminant function analysis results are reported using cross-validated classification rates with the inclusion of posterior probabilities for group membership. An ancestry bias estimate is also provided along with the results of those individuals correctly classified. Figure 21 provides a visual summary of the percentage of individuals correctly classified with the associated ancestry bias, which is the difference in correct assignment of sub-Saharan African and European individuals according to each predictor variable used, as well as the distributions of the posterior probabilities of correctly classified individuals in the discriminant analysis. Figure 22 provides another visual representation of the data but in table form. The figure includes the percentages of correctly classified individuals with the ancestry bias taken into account in the correct classification rates and a colour gradient representing high and low confidence in correct classifications (i.e. posterior probabilities). A complete output of discriminant function analyses including associated ancestry biases and lower, upper

and mean ranges of posterior probabilities can be found in the Appendix in Tables 1 to 6.

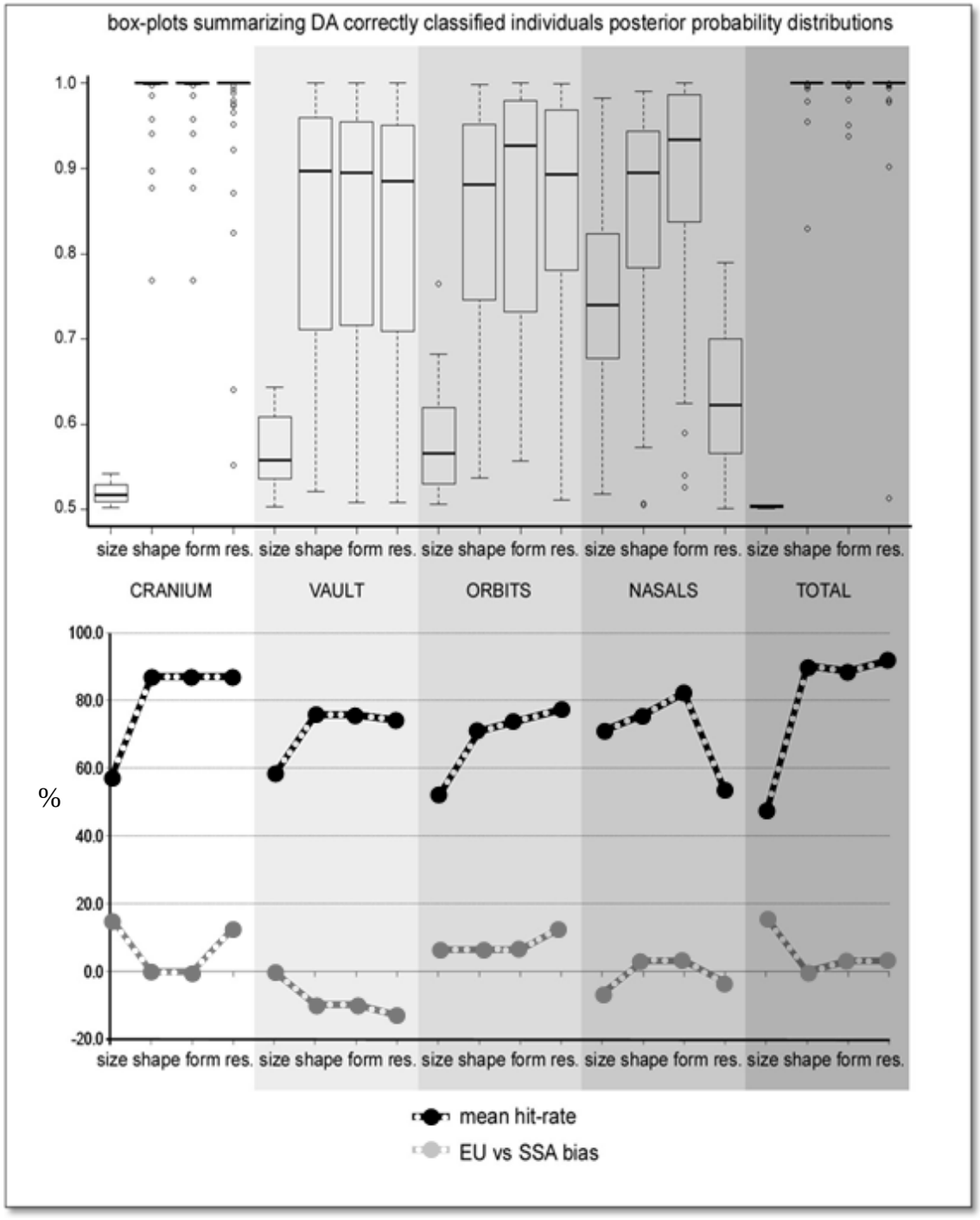

Figure 21 Summary of results obtained from the discriminant analyses (DA) using all datasets and all variables employed in the study. Correct classification results are shown along with associated ancestry biases and posterior probability distributions. 'Res.' stands for 'common allometric residuals (i.e. 'sized-corrected' form), and EU and SSA stand for European and sub-Saharan African individuals respectively. Image was created by Dr. Andrea Cardini, a co-supervisor on the project.

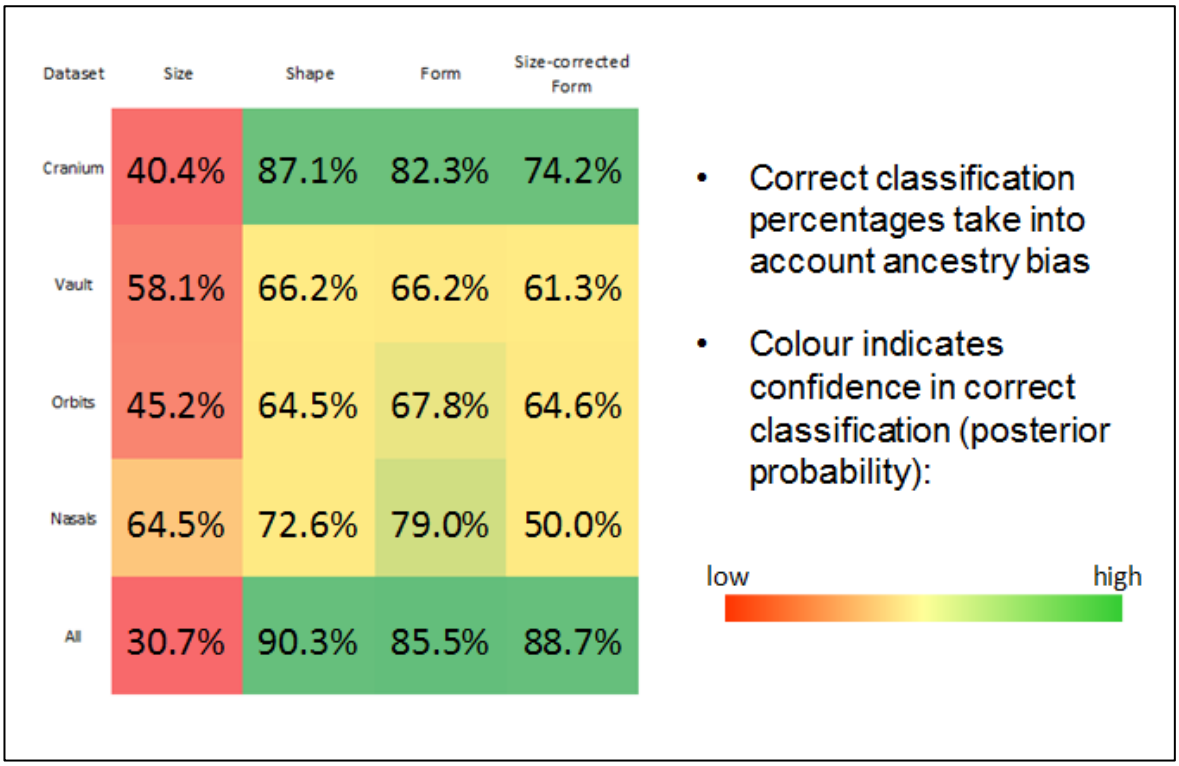

Figure 22 Summary of results obtained from discriminant analyses in the form of a table using all datasets and all variables employed in the study. Correct classification percentages are shown and associated ancestry biases are included in the correct classification rates. Posterior probabilities are also depicted in the form of a colour gradient where green represents high confidence and red low confidence. 'All' stands for all landmarks and semilandmarks employed.

Overall, using centroid size as a predictor for ancestry classification across all datasets correctly classified ancestry groups between 46.8-71.0% of the time with an overall mean of 56.8% correct classification. Posterior probabilities for correctly classified individuals using size were generally low with a mean range from 0.504-0.746. Only the semilandmarks on the lower nasal aperture were able to correctly classify ancestry with a fairly high accuracy (mean classification accuracy = 71.0%) and a certain degree of confidence (mean posterior probability = 0.746) using centroid size as the predictor variable, however, slightly more Europeans were correctly classified than sub-Saharan Africans as indicated by an associated ancestry bias calculation of -6.5%. In all other cases, size poorly classified individuals into their respective ancestry groups. The lowest classification rates for size occurred with the dataset using all

landmarks and semilandmarks. 46.8% of cross-validated grouped cases were correctly classified and the associated ancestry bias was 16.1%.

In general, shape confidently predicted ancestry across all categories and groups. Using shape as a predictor correctly classified between 71.0%-90.3% of individuals with an overall mean of 80.0% correct classification from all datasets. Posterior probabilities for correctly classified individuals were generally high with a mean range from 0.838-0.995. Ancestry bias values were low and ranged between no recorded ancestry bias to 3.2% from the standard landmark set, the lower nasal aperture and the entire landmark and semilandmark configuration. Correct classification rates for these datasets were 87.1%, 75.8%, and 90.3% respectively. The lowest value of correctly classified individuals using shape as a predictor occurred with the semilandmark dataset arranged along the orbital rim: 71.0% of grouped cases were correctly classified in terms of ancestry and the ancestry bias was 6.5%. Semilandmarks along the vault correctly classified 75.8% of individuals and the associated ancestry bias value was -9.6%.

Discriminant analyses on form correctly classified between 74.3%-88.7% of all individuals with an overall average of 81.3% for correct classification from each dataset. Mean values for posterior probabilities from each dataset ranged from 0.838-0.997. Ancestry biases were similar to the values associated with those from shape – the standard landmark set, the lower nasal aperture and the entire configuration of all standard landmarks and semilandmarks were less than or equal to 3.3%. As in the shape analysis, the semilandmarks along the vault and orbital rim had the lowest correct classification rates (75.8% and 74.3% respectively) and highest ancestry biases (-9.6% and 6.5% respectively).

‘Size-corrected’ form discriminant analyses correctly classified 53.2%-91.9% of all individuals into their respective ancestry groups with an average of 76.8% of individuals correctly classified from each dataset. The dataset that correctly classified the least number of individuals was the lower nasal aperture and the dataset that correctly classified the most individuals was the entire configuration of all cranial landmarks and semilandmarks. Average values for posterior probabilities from all datasets ranged from 0.838-0.995. Ancestry bias values were 12.9% for the standard landmark configuration and orbit dataset and -12.9% for the vault dataset. Ancestry biases for the lower nasal aperture and the entire configuration of all landmarks and semilandmarks were -3.2% and 3.2%, respectively.

‘Chance-corrected’ cross-validated correct classification rates using the TAU statistic for the various datasets and predictor variables are presented in Figure 23. ‘Chance-corrected’ classification rates for size ranged from -6.5% to16.1%; shape ranged from 41.9% to 80.6%; form ranged from 48.4% to 77.4%; and ‘size-corrected’ form ranged from 6.5% to 83.9%. Detailed results of ‘chance-corrected’ accuracy rates from all dataset using all variables can be found in the Appendix in Table 7. Using the standard landmarks dataset on shape as an example, 87.1% of individuals were correctly classified into their respective ancestry groups from the discriminant analysis and 74.2% of individuals were correctly classified when the effect of predicting group membership by chance alone is taken into account. For the lower nasal aperture semilandmark configuration with size as the predictor variable, the correct classification rate from the discriminant analysis was 71.0% and the ‘chance-corrected’ correct classification rate was 41.9% which is about 42% less errors than expected by chance.

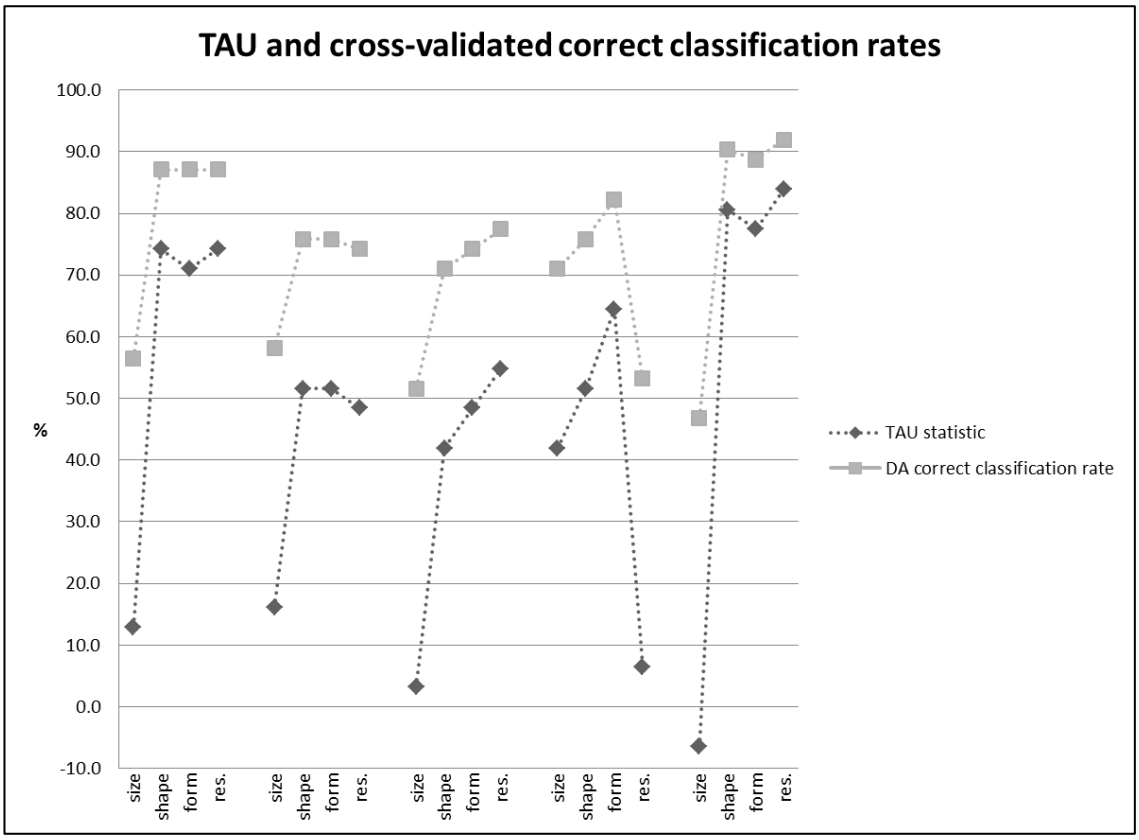

Figure 23 Classification accuracy rates of the discriminant analyses and the TAU statistic expressed as a percentage.

Typicality probabilities for the cranial landmark configuration from the discriminant function analysis on shape ranged from 0 to 0.910. A total of 27 individuals had typicality probabilities less than 0.001. This suggests that the morphology of the crania from these individuals may not be typical of the groups into which they were assigned. Figure 24 shows the scatterplot of the first two principal components of shape from the cranial landmark dataset with individuals with typicality probabilities less than 0.001. Fourteen European individuals and thirteen sub-Saharan African individuals had typicality probabilities less than 0.001. In general, the scatterplot of the first two PCs of shape show that sub-Saharan African individuals with low typicality values appear as outliers or tend to plot along the margin of the European

cluster whereas the European individuals with low typicalities do not seem to plot in a characteristic or identifiable way; they plot within the European cluster with some individuals plotted as outliers. Typicality probabilities from the discriminant analysis on size using the lower nasal aperture dataset were all above 0.001 and ranged from 0.003 to 0.981.

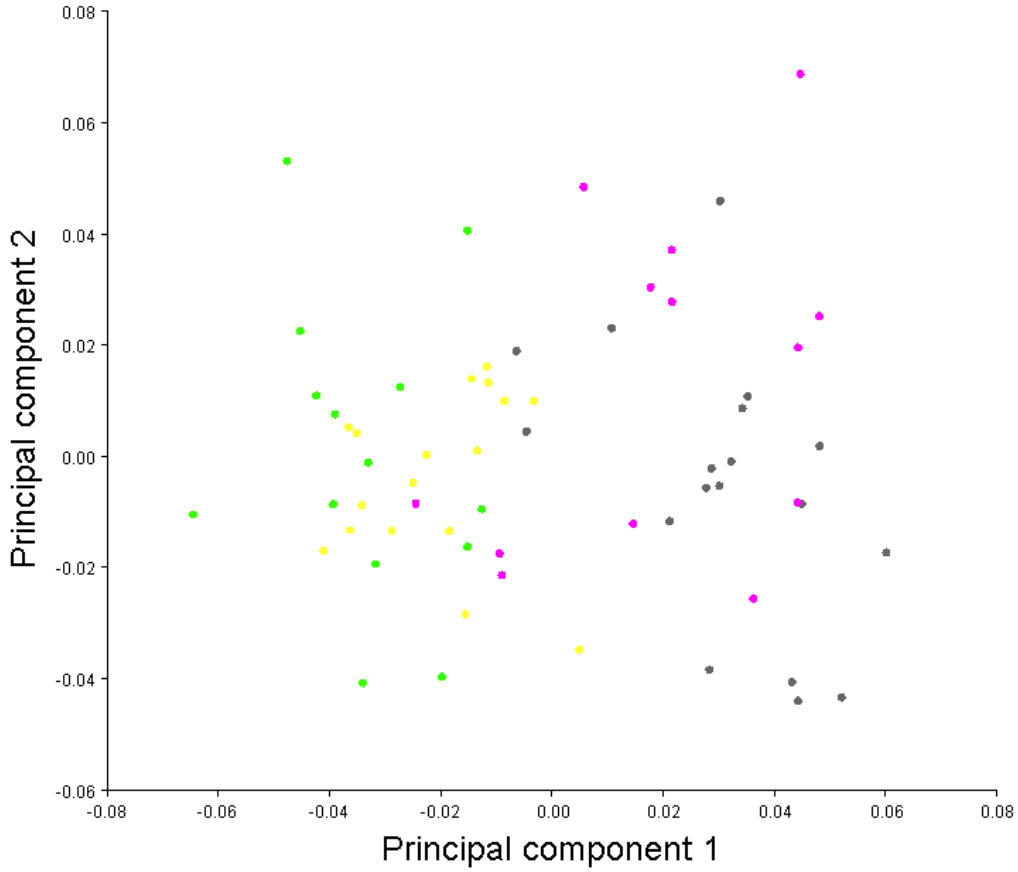

Figure 24 Scatterplot of first two principal components of shape from the cranial landmarks. Pink and yellow dots represent sub-Saharan African and European individuals, respectively, with typicality probabilities less than 0.001. Dark grey dots are sub-Saharan African individuals with typicality probabilities greater or equal to 0.001 and light yellow dots represent European individuals with similar typicality probabilities.

5. DISCUSSION

The present research project was carried out to explore shape, size, and form differences between crania from individuals of sub-Saharan African and European ancestry using 3D geometric morphometrics and discriminant function analyses with leave-one-out cross validation. The specific aim of the study was to determine if it was possible to classify individuals accurately into their respective geographic ancestry groups based on configurations of landmarks and/or semilandmarks. Classification procedures were carried out using the entire cranial landmark/semilandmark dataset, subsets of the data focused on specific features of the cranium, and a standard cranial landmark set.

Despite the small sample size, the results of the study indicate that 3D geometric morphometric methods conducted on 3D images of the cranium are able to classify geographic ancestry with a significant level of consistency and confidence. The degree of success of the classification, however, was dependent on the landmark/semilandmark configurations and predictor variables used.

In evaluating the effectiveness of the various three-dimensional multivariate descriptors to classify ancestry, the correct classification rates, the posterior probabilities and the ancestry bias calculations must all be taken into consideration to determine the best predictor. The TAU statistic should also be examined to provide details that explain the improvement of classification accuracy over what would be expected by random assignment. Although the percentage of correct classification achieved from the discriminant analysis function is the most important measure of discrimination, TAU provides a means of estimating the magnitude of the correct classification percentage relative to the percent of correct classifications made by

random chance assignment. An ideal dataset should have a high cross-validated accuracy rate (at least above 80%), high posterior probabilities (as close to 1.0 as possible), and a low ancestry bias set at less than 5%, as proposed by Franklin and colleagues (2012), in order to be considered suitable for application in forensic human investigation.

The standard cranial landmarks and the entire landmark and semilandmark configuration using all variables, excluding size, classified individuals with the highest accuracy and confidence rates (i.e. posterior probabilities) and relatively low associated ancestry bias calculations (all less than 5% except for the standard landmark set using `size-corrected` form as the variable). Accuracy rates for the standard landmark dataset using shape and form as variables and the total landmark/semilandmark configuration using shape, form and 'size-corrected' form variables surpassed 80% correct classification rates even when the ancestry bias calculations were taken into consideration and subtracted from the accuracy rates. Actual accuracy rates were higher and ranged from 85.5%-91.9% with high overall posterior probabilities. There was very little variation in terms of the entire cranium that could be attributable to size differences between the groups since correct classification rates and posterior probabilities were low and associated ancestry biases were high for the standard landmark dataset and entire landmark/semilandmark dataset for this variable.

When size was added to shape for the analyses on form and 'size-corrected' form across all the datasets, the effect on classification accuracy was inconsistent: in some cases there was slight improvement; in other cases there was no change at all and in still other cases the classification rate was lower than using shape analysis alone. This clearly indicates that shape, and not size, is the more important characteristic in

analyzing differences between crania in terms of ancestry. As long as variables of shape are in the ancestry classification analyses for all datasets, especially using the standard landmarks and the entire configuration of landmarks/semilandmarks, there is an overall improvement in classifications accuracy rates.

In terms of the specific features analyzed, the nasal aperture was, in general, the more accurate in distinguishing between the two ancestry groups using all of the variables except for 'size-corrected' form. This is not surprising since the nasal aperture region is often considered one of the most useful craniofacial traits for assessing ancestry using traditional morphological and metric methods (Gill, 1998). Semilandmarks along the lower nasal aperture performed better using the variable of size alone in classifying individuals into their respective ancestry groups in comparison to all other datasets. The accuracy rate for the lower nasal aperture using size was 71% with an ancestry bias of 6.5%, and mean posterior probability of 0.746. All other landmark and/or semilandmark configurations performed significantly worse when size was the only variable analyzed. For these other configurations, correct classification rates ranged from 46.8% to 58.1% with very low posterior probabilities for those correctly classified and relatively high ancestry biases. Similarly, as with all the other datasets, semilandmarks along the lower nasal aperture performed better when shape was included in the analyses except with the variable of 'size corrected' form. Correct classification rates were 72.6% using shape alone and 79.0% using form when the differences between accuracy rates and ancestry biases were included. The fact that the nasal aperture set performed so poorly with the 'size-corrected' form variable, which controlled for the effect of allometry using a common size to predict shape for analyses of 'non-allometric' variation regardless of group, suggests that differences in shape for

this feature between the groups are somewhat dependent on size variation (Viscosi and Cardini, 2011).

Size differences between the groups using the lower nasal aperture is in agreement with previous studies which demonstrated that individuals with origins from Europe tend to have a narrower nasal aperture than individuals with ancestry origins from sub-Sahara Africa (Gill, 1998). The visualizations of the shape changes along the first PC in the direction of the sub-Saharan Africans and Europeans as well as the images of the mean forms from either group (shown in the results section in Figure 19 and Figure 20), illustrate this difference between the ancestry groups well. A linear measurement such as nasal breadth, which is commonly applied in traditional metric ancestry methods, measured from the bilateral anatomical landmark alare, should provide a quick and simple means to discriminate between the groups, albeit not with the highest confidence and accuracy rates, at least based on the sample of individuals from this project. The typicality probabilities (all greater than 0.001) calculated for this dataset and the relatively high posterior probabilities demonstrate that the variable of size can be used fairly confidently in predicting ancestry. The fact that individuals were classified with higher accuracies, higher confidence, and ancestry biases less than 5% using shape and form variables again reiterates that shape is an important discriminating factor for ancestry prediction. When size was added to shape for the form analysis accuracy rates achieved their highest correct classification indicating that both size and shape of the lower nasal aperture are significantly different between individuals with European and sub-Saharan African ancestry origins.

Classification accuracy based on the vault and orbits for the different variables was not as high as those for the other datasets. Analyses on shape, form, and 'size-

corrected' form for the orbital and vault datasets had accuracy rates which ranged from 61.3% to 67.8% when the differences between correct classification rates and the associated ancestry bias were included in the classification accuracies. The lowest accuracy rates for both these datasets occurred when size was the only variable used in the discriminant analysis. The highest accuracy rates, with the inclusion of associated ancestry biases, occurred with shape and form variables using the vault dataset, and with form variables using the orbit dataset. Orbital and vault shape are both considered to differ between the two ancestry groups in traditional morphological and metric analyses but the discriminant analyses using these variables did not classify individuals with high confidence or accuracy, and ancestry biases were all well above 5%. Morphologically, the orbital shapes of individuals with sub-Saharan African origins tend to appear rectangular and those with European ancestry tend to appear more angular, sloping downwards with a rhomboid-like shape. Traditional metric evaluations of the orbits for ancestry classification involve calculating the height of the orbit, dividing by the breadth and multiplying the computed value by 100 (Husmann and Samson, 2011). When this orbital index is used as part of metric ancestry assessment, European and African individuals tend to fall within the intermediate (mesoseme) range of variation (Husmann and Samson, 2011; Ukoha et al., 2011). The orbital index conveys very little shape information and is basically a rectangle: whether the rectangle is longer in the up-down direction, longer in the right-left direction, or square. This index would not be able to differentiate a rectangular from a rhomboid-like shape. Perhaps one could argue that the differences associated with the size and shape of the orbits from these two ancestry groups is not strong since both morphologically and metrically the differences may not be compelling: visually they are both seen as angular and unrounded between groups, and rhomboidal and rectangular shape differences may

be small; and metrically shape information is not conveyed and individuals from either ancestry group tend to fall within the same intermediate category. The differences, hence, may be subtle and this could be a reason for the less successful differentiation rates obtained from the orbital dataset. However, this could also be an effect of the small sample size employed in the research and perhaps a larger sample would increase classification accuracy using this feature. As was mentioned in the literature review section, Husmann and Samson carried out a study applying 2D geometric morphometric methods to study the shape of the orbit for sex and race estimation using a sample of 184 Black women, 236 Black men, 110 White women, and 232 White men. They found that orbital shape between Blacks and Whites demonstrated little shape difference. The percentage of attributable shape variation was actually quite small (1.36% for race differences) and principal component scatterplots almost completely overlapped between the groups using this feature (Husmann and Samson, 2011). Mean shapes of the orbits between all females, all males, Blacks (both sexes), and Whites (both sexes) were presented and differences between groups were minimal. It is possible that landmarks on 2D images of the orbit do not capture the morphological complexity of the curved surface of this feature. The results obtained in this research were below the 80% accuracy limit for the orbital aperture and the principal component scatterplot of the first two PCs of shape (shown in the results section in Figure 16) displayed nearly complete overlap. However, the shape and form of this feature should not be ruled out as a potential trait to aid in ancestry identification.

As for the semilandmarks along the vault in the midplane, the differences between the two ancestry groups are, again, likely to be more subtle than the differences usually observed from such areas as the nasal aperture. Ideally, the semilandmarks along the vault on the midplane would have picked up the morphological differences

frequently observed between the two ancestry groups. However, classification rates were also well below 80% and ranged from 58.1% to 66.2% using all variables if ancestry biases were included in the correct classifications rates. European crania tend to have a mesocephalic head shape with a sloping upright profile and a pronounced external occipital protuberance and sub-Saharan African crania tend to have a dolicocephalic head shape with a sloping but more rounded frontal profile and a post-bregmatic depression (Randolph-Quinney et al., 2009; Wilkinson, 2004). In traditional metric analyses head length can be measured with calipers using certain anatomical landmarks but features such as a post-bregmatic depression or external occipital protuberance cannot be easily quantified. Geometric morphometrics provide a means by which these types of features can be quantitatively analyzed; however, differentiation between the two sample groups employed in this research was not satisfactorily apparent using this feature alone.

The 'chance-corrected' accuracy rates expressed by the TAU statistic provide a measure of how well a technique performs if classification by pure chance alone is removed from the correct classification rate. High rates of TAU indicate that the landmarks/semilandmarks and variables employed are reliable and effective in discriminating between the groups beyond chance classifications. Any TAU value greater than zero indicates an improvement in the classification process beyond mere chance and the higher the TAU value the greater the improvement. All datasets, except for the total configuration of landmarks/semilandmarks using the variable of size, correctly classified individuals with an accuracy rate above chance. The discriminant analysis using the entire configuration of landmarks/semilandmarks using size classified individuals with a low success rate and the 'chance-corrected' classification rates revealed that those individuals correctly classified were categorized into respective

groups with a success rate worse than if individuals were randomly assigned to groups. This reiterates that the variable of size should not be utilized to differentiate between European and sub-Saharan Africans from features over the entire cranium, at least for the sample of individuals utilized in this research. The highest 'chance-corrected' accuracy rate (i.e. the TAU value) occurred with the total configuration using the variable of 'size-corrected' form and was calculated as 83.9% .This indicates that the discriminant analysis classification made approximately 84% fewer errors than would have been expected on the basis of random assignment and that the technique is reliable at classifying individuals. In other words, since two category groups were used in this study, individuals had a 50% chance of being wrongly (or correctly) classified based on random assignment, and of the 50% that would have been wrongly classified by chance alone, approximately 84% of these individuals were correctly classified by the discriminant function analysis classification. All datasets using shape and form using 'chance-corrected' classification rates accurately classified individuals above at least 40% (i.e. 40% fewer errors than random assignment alone). The standard landmark and total landmark/semilandmark configurations using shape and form and 'size-corrected' form were by far the highest for 'chance-corrected' classification rates and were all above 71.0% and ranged up to 83.9% indicating that classification based on these dimensions is good over a random allocation of individuals into groups.

The standard landmark set and entire landmark/semilandmark configuration (i.e. landmarks captured over the entire cranium) were significantly better at classifying ancestry than any subsets of points digitized along specific cranial features, and the most successful predictor variables were those that included shape in the analyses. For single traits, the nasal aperture dataset, using the variable of form, approached 80% accuracy with the ancestry bias included in the correct classification rate. This trait

should be considered useful for ancestry prediction between European and sub-Saharan African individuals. This feature could be considered particularly important when incomplete or damaged cranial remains are found but the nasal aperture area is intact. An overall aspect of the cranium with single landmarks placed more loosely on multiple traits commonly used in ancestry identification (i.e. the standard landmark set) appears to provide better ancestry classification than more densely gathered points on specific ancestry-related features. This is not surprising since it is frequently recommended that multiple traits should be examined to identify ancestry from skeletal remains to attain maximum reliability instead of relying on single traits. However, even though some of the specific features, such as the orbits and vault, exhibited only mediocre classification success in this study, their potential importance in assessing ancestry should not be overlooked and further studies using larger sample sizes should be carried out. Human remains are frequently found damaged and incomplete and the use of single features to assess ancestry is sometimes all a practitioner can rely on.

In terms of the best model for ancestry classification purposes, the most accurate descriptors were the entire configuration of landmarks and semilandmarks using shape, form and 'size corrected' form as the variables (accuracy rates ranged from 88.7% to 91.9% with ancestry bias less than or equal to 3.2% and a mean posterior probabilities of greater than or equal to 0.995). The standard landmark set using the variable of shape and form, however, is likely to be the preferred model for ancestry classification since it is based on fewer variables (sixty-eight standard landmarks across the cranium as opposed to one hundred forty-eight landmarks/semilandmarks from the entire configuration). The posterior probabilities and ancestry bias were generally very similar to the entire landmark/semilandmark configuration using the variables of shape and form. The accuracy rates for the standard landmark set using the variable of 'size-

corrected' form were high, however, the associated ancestry bias was 12.9% and thus it is not included. In terms of the time and ease of acquisition the standard landmark dataset would likely be more desirable. However, the typicality probabilities that were calculated for the standard cranial landmark dataset were surprisingly low. Despite having high posterior probabilities this suggests that some individuals may not have been typical of any of the groups used in the analyses even though the individuals correctly classified were closer to the group to which they were assigned. The low typicalities for some individuals may also be an artifact of the heterogeneous nature (since sex and ethnicity/nationality were not taken into consideration in analyses) and temporal inconsistency of the sample, however a study using a significantly larger sample size would be required to investigate this effect more definitely.

The reliability of ancestry identification using the non-metric approach is unclear and difficult to substantiate although some practitioners claim ancestry prediction accuracy rates of up to 90% (Clement and Ronson, 1998). Skilled practitioners of the non-metric approach certainly use identifiable cranial features in their classification criteria but they may also be prone to subjective bias and intuition. With traditional metric methods, using craniometric discriminant functions and the computer program Fordisc, accuracy rates have been reported, in some studies, to approach the accuracy rates achieved in this research using the standard landmark dataset (up to or greater than 80%) but have proven to have inconsistent results in other studies. It is important that methods be developed that can be statistically quantifiable as well as retain the complex morphology of the skull. This research and the results obtained should be seen as preliminary and exploratory, however, this study demonstrates that it is possible to quantify ancestry variation with accuracy while still retaining the complex morphology of the skull. The classification accuracy rates

associated with the standard cranial landmarks and total configuration of landmarks/semilandmarks with all variables using shape as components indicate 3D geometric morphometrics to be a promising method for identifying ancestry from unknown individuals. It appears that ancestral variation can be measured and quantified at least between the groups used in this study. Differentiation and classification between ancestry groups using geometric morphometrics methods should continue to be researched using larger samples and multiple ancestry groups in order to gain further insight into the power of this method for ancestry identification purposes.

There are not a lot of other studies that have been conducted involving geometric morphometrics and classification of ancestry for human identification using the cranium. The principal goals of some of the studies involving ancestry, the cranium, and geometric morphometrics were exploratory in nature and generally concerned assessing overall facial shape differences between various geographic populations (Viðarsdóttir et al., 2001; Henessy and Stringer, 2002). Results of these types of studies showed that geometric morphometrics could characterize shape accurately and that variations between groups could be visualized clearly and defined confidently in a statistical sense. Although classification accuracy was not the main objective of the study, Viðarsdóttir and colleagues were able to classify correctly between 66.7% and 100% (mean 82.6%) of adult individuals using cross validation from 10 different geographic populations using 26 unilateral landmarks on the cranium. The main objective of their study involved using geometric morphometric to determine population differences in the growth and development of the human facial skeleton. They were able to correctly classify 75% of African American adults and 77.78% of Caucasian adults from their sample. Sample size, however, was limited and group sizes were unbalanced: the largest group of adults was represented by African Americans and was comprised of

only twelve individuals and the smallest group was represented by Polynesians and comprised of five individuals. When sub-adults were included in analyses along with adults, classification accuracies fell to between 53% to 88% with an average of 71% of all individuals correctly assigned. Average mean correct classification rates for allocation of adult crania to populations were high, however, the unequal sample sizes among groups and the small sample are likely to increase over-fitting of individuals in the discriminant analysis and reduce generalizability of results (Kovarovic et al., 2011).

Ross and colleagues (1999), in their study involving geometric morphometrics and allocation of crania to American Black and White groups using variables of shape, achieved results which corresponded closely to those using traditional approaches. The accuracy rate achieved on a sample of 19 American Blacks and 19 American Whites using fourteen superimposed homologous cranial landmarks was 84.2% using cross validation. The authors also compared this classification rate to discriminant analyses based on traditional linear measurements from the cranium and achieved an overall accuracy of 78.9%. Buck and Viðarsdóttir (2004) used geometric morphometrics to identify race from sub-adult mandibles. They employed a large sample, consisting of 174 mandibles, from five ancestral populations (African Americans, Native Americans, Caucasians, Inuit, and Pacific Islanders) using 17 unilateral landmarks on the mandible and achieved an average of 70.1% correct classification. They then used a reduced sample consisting of three of the populations (African American, Caucasian, and Native American) and correctly classified 87.6% of individuals into their respective ancestry groups.

Although the program 3D-ID has been created and is in use to classify individuals by ancestry using 3D geometric morphometrics, there is not yet any

information available in the peer-reviewed literature relating to the accuracy rates associated with this software. So, it is not yet possible to compare the results obtained in this analysis to those obtained with 3D-ID. In this analysis, however, more standard landmarks and more aspects of the cranium related to ancestry variation were utilized in comparison to 3D-ID. In addition, in the present research, specific features on the cranium were examined with semilandmarks to determine their possible utility in ancestry identification as standalone features. Semilandmarks are not used in the 3D-ID program.

5.1. LIMITATIONS

There are limitations involved in this study which may have had some effects on the results obtained. Most of the limitations concern the nature and size of the sample of individuals used in the study although some relate to intrinsic factors that may be more difficult to control.

5.1.1. INHERENT ERRORS

Pertaining to the 3D images employed in the study, two different types of data were used: images from a surface scanner (48 images in total) and images from CT scans (14 images in total). No study was conducted to compare the different types of scanning technology. Scanning equipment may differ in volume and resolution, however, the spatial resolution of 3D images produced by laser scanning devices is generally considered equal to or greater than CT images (Fourie et al., 2011). Fourie and colleagues conducted a study in which they evaluated the accuracy and reliability of different 3D scanning systems, including laser scanning equipment and CT data. They found anthropometric measurements on images created using different 3D imaging technologies were comparable in accuracy to direct measurements taken with callipers.

They also found that the measurements obtained from 3D images created using different imaging technologies were reliable and accurate between instruments and that the measurements obtained from the different systems could be combined for research purposes (Fourie et al., 2011). Multiple other studies have also shown that direct anatomical measurements are highly correlated with anatomical measurements derived from CT and surface scan data (Waitzman et al., 1992; Aung et al., 1995; Richtsmeier et al., 1995; Park et al., 2006; Dean et al., 2009).

As was discussed previously, there are three types of landmarks and the accuracy with which they can be located is subject to varying degrees of error. The relative locations of Type II and Type III landmarks are generally considered more ambiguous than Type I landmarks. Although Type I landmarks are considered the most accurate, there are few points on the cranium that can defined according to their strict definition. Even though the locations of Type II and III landmarks are less clear and may be more prone to inter- and intra-observer errors, the use of these types of landmarks can enhance interpretation of shape changes since it allows for more points to be captured on specimens and potentially more biologically meaningful results to be obtained.

In the opinion of the author, the ability to locate certain anatomical landmarks differs considerably between the two types of images used. Type I landmarks, for example, were generally difficult to locate on surface scan images since discoloured or complex surface morphology was not always picked up by the laser scanning device. Locating these types of landmarks on CT scans, however, was generally more reliable since this method provides more surface detail. Precise locations of anatomical landmarks on the surface scans, therefore, may not correspond completely to their exact

positions on the dry skull. The discrepancy in landmark positioning on the surface scans utilized in this project has been demonstrated to be negligible in terms of intra-observer error since both types of 3D images were used in the digitization error study. For a follow-up study an inter-observer error test should be carried out to determine whether the landmarks captured on 3D cranial images are precise and repeatable between observers. A study to determine the precision of locating individual landmarks, instead of landmark precision as a whole, on 3D images of the cranium would also be an important consideration.

5.1.2. SAMPLE BIAS

In this study the sample size was relatively small and some of the individuals may not have been reflective of the ancestry group to which they were assigned. In other words, the group of individuals in this sample may not have optimally reflected biological reality. Individual variation may have influenced classification results since traits associated with ancestry prediction are not absolute or discrete. Human variation is complex and continuous and the individuals within the sample may or may not represent the variation typically observed within their populations.

The sample also included crania from varying temporal periods. This may affect the study's ability to classify accurately individuals in a modern context. Almost all of the sub-Saharan African crania were historical in origin, dating to around the nineteenth century, and there may be a bias associated with the classification accuracy of these individuals as a result. It has been demonstrated that biological secular change in humans has taken place within as little as 100 years and that significant changes in the shape and form of the skull have been observed between individuals from both ancestry groups employed in this research (Gravlee et al., 2003; Wescott and Jantz, 2005). Of

possible relevance, however, is a study carried out by Jantz and Wescott using geometric morphometrics to assess craniofacial secular change in American Blacks and Whites. The study showed that secular changes to the superior vault and face in both ancestry groups was minimal although there was some change observed in the base of the cranium (Wescott and Jantz, 2005). Most of the landmarks and semilandmarks employed in this study were concentrated on the craniofacial region and may have been relatively unaffected by temporal change. Ideally, however, this study should be replicated on a larger sample of modern crania to provide a reasonably unbiased representation of modern day cranial variation between the two groups.

Individuals in this study were not grouped by sex in the analyses since, for most of the sample, this information was not available. Sex was treated as a source of 'noise' which may have affected correct classification rates and resulted in more difficulty discriminating between the groups. Sex and ancestry traits are interconnected and the sex of an individual can affect the way ancestry traits are perceived and vice versa. For example, a moderate to heavy supraorbital ridge and a large and long mastoid process is associated with Caucasoid ancestry. However, these traits are also typical of general male morphology. Pooling of individuals by sex could improve ancestry classification accuracy since female and male traits would have been somewhat controlled.

The crania employed in this research also represent individuals from a number of different nationalities and ethnicities. Regional variation on a smaller scale occurs between populations and this may have affected classification accuracy, and may be partly responsible for the very low typicality probabilities computed for shape from the standard landmark dataset. However, classification accuracy was primarily based on the broad labels used in forensic anthropology for ancestry identification of unknown

individuals and not taking smaller scale regional variation into consideration may be appropriate for the direction of this research.

5.2. IMPLICATIONS FOR FURTHER RESEARCH AND STUDY

In summary, this study has demonstrated that the use of geometric morphometric methods has the ability to classify ancestry accurately between individuals of sub-Saharan and European ancestry using landmarks that cover the entire cranium with shape and form as variables. The work has also demonstrated that semilandmarks along the lower nasal aperture using the variable of form can classify individuals with an accuracy rate close to 80%.

Geometric morphometric techniques offer some significant improvements over traditional approaches to estimate ancestry:

1. Geometric morphometric techniques are more objective than the non-metric approach since the method relies on anatomical landmarks that are well-defined in the literature (Gonzales et al., 2011)
2. The method allows for the quantification of the complex geometry of the cranium and traits on the cranium to be better retained and described than traditional metric methods
3. The method allows pure shape information to be easily singled out from size
4. The method allows for differences in shape and form between individuals and groups to be intuitively visualized in principal component shape space

Landmarks and semilandmarks were shown to effectively portray the cranial *gestalt* and other features from the cranium as well as compare objectively size, shape, and form differences between crania from the ancestry groups studied. Although some specific features did not fare as well in ancestry predictions they should not be ruled out as potential descriptors. Further research needs to be conducted to establish if the methods employed and classification accuracies achieved in this study are in fact biologically meaningful. Although the results of this study are certainly encouraging, to make any definitive statements about the effectiveness and accuracy of the geometric morphometric method as a technique that could be applied in forensic identifications, a study using a larger sample size, as well as more robust controls (sex, regional/ethnic origin) is required. Future research directions should include the use of more ancestry groups of known sex individuals to develop discriminant functions which can both classify geographic origin as well as sex in adults using the cranium. Subsets of the standard landmark configuration should also be investigated in more detail to determine where most of the ancestry-related variation is actually occurring. Applying EDMA may also be valuable since this technique can provide more specific information about which landmarks differ the most between individuals and groups to isolate areas of greatest shape change. Traits semilandmarked could also be examined in combination (i.e. investigate whether using the orbit and nasal aperture datasets, for example, provide greater discrimination between groups if they are merged together). As analysis and visualization software is improved, techniques of 'sliding' the semilandmarks will also be refined which may improve classification accuracy as well as enhance the visualization of 3D semilandmark data in an intuitive way.

6. CONCLUSION

Presently, there is little research being conducted to evaluate or enhance traditional ancestry classification methods (Christensen and Crowder, 2009). The assessment of ancestry/race plays an important role in the identification of human remains in a forensic context and there is a need to improve and quantify methods to predict ancestry if this component is to be admissible in a court of law and continue to be used in forensic human identification. Out of all the components of a forensic anthropology biological profile, estimation of ancestry is considered the most difficult. Traditional methods used to predict ancestry may be problematic and the traits used to distinguish between groups do not point directly to one population over another. The traits are numerous, not 'fixed', have different frequency of occurrence rates, and are subject to individual variability. Determination of ancestry from skeletal remains in a forensic setting using traditional techniques may be problematic especially if considered from a judiciary perspective "since they employ a combination of traditional scientific methodologies and less rigorous observational methodologies coupled with case study evaluations or casework experience" (Christensen and Crowder, 2009, 1211).

The analysis of crania from 3D images using geometric morphometric techniques provides an alternative technique to identify and classify ancestry from unknown skeletal remains in a repeatable, accurate, and quantifiable manner. The use of 3D imagery is becoming increasingly more common in physical anthropology and studies employing this kind of data provide confirmation and validation of the use of scanning technology in morphometric analyses. The present study has shown that ancestry can be accurately predicted between European and sub-Saharan African groups using virtual images and variables of shape and form (i.e. size and shape) when

landmarks and/or semilandmarks are captured over the cranial *gestalt*. The subsets of semilandmarks using all variables (size, shape, form, and 'size-corrected' form) generally did not classify ancestry between the groups with a high degree of accuracy. However, semilandmarks along the lower nasal aperture approached high levels of accuracy and this trait should be considered an important sub-region for ancestry classification. Size differences alone between crania and features on the cranium did not classify ancestry accurately or confidently between the groups employed in this research. This implies that linear distance measurements as used in traditional metric ancestry assessments, which rely predominantly on length and width measurements of anatomical landmarks, may not be the most optimal method to differentiate between ancestry groups. Shape has been shown in this project to be more important in classifying ancestry than size differences alone.

As an extension of the research carried out here, more effective discriminant functions could be developed for ancestry classification using cranial data. The results obtained in this study as well as the data collected could be utilized in the creation of classification software or in the enhancement of software already available (e.g. 3D-ID) to identify ancestry from crania in a forensic context. Geometric morphometrics may never replace entirely traditional methods for ancestry identification but it can provide quantitative shape information that can corroborate and validate methods already in use. Geometric morphometrics will undoubtedly provide an increasingly powerful and valuable tool in assessing shape and shape change among and between individuals and groups and has the potential to objectively resolve some of the problems associated with ancestry identification in forensic anthropology.

LITERATURE CITED

Adams, D. C., Rohlf, F. J., and Slice, D. E. (2004). Geometric morphometrics: ten years of progress following the 'revolution'. *Italian Journal of Zoology*, 71:5-16.

Albanese, J. and Saunders S. R. (2006). Is it possible to escape racial typology in forensic identification? In: Shmitt, A., Cunha, E. and J. Pinheiro (Eds.). *Forensic Anthropology and Medicine: Complementary Sciences from Recovery to Cause of Death*. Totwa: Humana Press Inc.

American Association of Physical Anthropologists (AAPA). (1996). AAPA statement on biological aspects of race. American Journal of Physical Anthropology, 101:569-570.

Aung, S. C., Ngim, R. C. K., and Lee, S. T. (1995). Evaluation of the laser scanner as a surface measuring tool and its accuracy compared with direct facial anthropometric measurements. *British Journal of Plastic Surgery*, 48(8):551-558.

Bass W. M. (2005). *Human Osteology*, 5th ed. Columbia, Missouri: Missouri Archaeological Society.

Black, S. (2000). Forensic osteology in the United Kingdom. In Human Osteology in Archaeology and Forensic Science. Margaret Cox and Simone Mays (Eds.) London: Greenwich Medical Media Ltd.

Bolnick, D. A. (2008). Individual ancestry inference and the reification of race as a biological phenomenon. In: B. A. Koenig, S. S.-J. Lee, & S. S. Richardson (Eds.) *Revisiting race in a genomic age*. New Brunswick, NJl: Rutgers University Press.

Bookstein, F. L. (1982). Foundations of morphometrics. *Annual Review of Ecology and Systematics*, 13(1): 451-470.

Bookstein, F. L. (1997). Landmark methods for forms without landmarks: Morphometrics of group differences in outline shape. *Medical Image Analysis* 1:225-243.

Bookstein F. L., Slice D., Gunz P., and Mitteroecker P. (2004). Anthropology takes control of morphometrics. *Collegium Antropologicum* 28, (Suppl. 2):121–132

Braga, J. and Treil, J. (2007). Estimation of pediatric skeletal age using geometric morphometrics and three-dimensional cranial size changes. *International Journal of Legal Medicine*, 121:439-443.

Bruner, E. and Manzi, G. (2007). Landmark-based shape analysis of the archaic Homo calvarium from Ceprano (Italy). *American Journal of Physical Anthropology*, 132: 355–366.

Buck, T. J., and Vidarsdottir, U. S. (2004). A proposed method for the identification of race in subadult skeletons. *Journal of Forensic Sciences*, 49(6):1-6.

Byers S. N. (2005). *Introduction to forensic anthropology*. 2nd ed. Boston: Pearson Education, Inc.

Bytheway, J. A. and Ross, A. H. (2010). A geometric morphometric approach to sex determination of the human adult os coxa. *Journal of Forensic Sciences*, 55: 859–864.

Cardini, A. (2012) Geometric morphometrics. *UNESCO Encyclopedia of Life Support*. In press.

Carpenter, J. C. (1976). A comparative study of metric and non-metric traits in a series of modern crania. *American Journal of Physical Anthropology*, 45: 337–343

Cartmill, M. (1998). The status of the race concept in physical anthropology. *American Anthropologist*, 100:651–660.

Caspari, R. (2003). From types to populations: a century of race, physical anthropology, and the American Anthropological Association. *American Anthropologist*, 105: 65–76.

Christensen, A. M. (2004). The impact of *Daubert*: implications for testimony and research in forensic anthropology (and the use of frontal sinuses in personal identification). *Journal of Forensic Sciences*, 49(3):427-430.

Christensen, A. M. and Crowder, C. M. (2009). Evidentiary standards for forensic anthropology. *Journal of Forensic Sciences*, 54: 1211–1216.

Clement, J. G. and Ronson, D. L. (1998). *Craniofacial Identification in Forensic Medicine*, London, Oxford University Press.

Cole, T. M. and Richtsmeier, J. T. (1998). A simple method for visualization of influential landmarks when using Euclidean distance matrix analysis. *American Journal of Physical Anthropology*, 107:273-283.

Curran, B.K. (1990). The application of measures of midfacial projection for racial classification. In *Skeletal Attribution of Race*. G. W. Gill and S. Rhine (Eds.). Maxwell Museum of Anthropology, Anthropology Papers No. 4: 53-57.

Dean, C. L., Lee, M. J., Robbin, M., and Cassinelli, E. H. (2009). Correlation between computed tomography measurements and direct anatomic measurements of the axis for consideration of C2 laminar screw placement. *The Spine Journal*, 9(3):258-262.

Dirkmaat, D. C., Cabo, L. L., Ousley, S. D. and Symes, S. A. (2008). New perspectives in forensic anthropology. *American Journal of Physical Anthropology*, 137: 33–52.

Elliott, M., Collard, M., (2009). FORDISC and the determination of ancestry from cranial measurements. *Biology Letters*, 5: 849-852.

Ferrario V. F., Sforza C., Pizzini G., Vogel, G., and Miani, A. (1993). Sexual dimorphism in the human face assessed by euclidean distance matrix analysis. *Journal of Anatomy*. 183:593–600.

Fourie, Z., Damstra, J., Gerrits, P. O., and Ren, Y. (2011). Evaluation of anthropometric accuracy and reliability using different three-dimensional scanning systems. *Forensic Science International*. 207(1-3):127-134.

Franklin, D., Oxnard, C. E., O'Higgins, P. and Dadour, I. (2007). Sexual dimorphism in the subadult mandible: quantification using geometric morphometrics. *Journal of Forensic Sciences*, 52: 6–10.

Franklin, D., Flavel, A., Kuliukas, A., Cardini, A., Marks, M. K., Oxnard, C., and O`Higgins, P. (2012). Estimation of sex from sternal measurements in a Western Australian population. *Forensic Science International*, 217: 230.e1-230.e5.

Galdames, I. C. S., Matamala, D. A. Z., and Smith, R. L. (2008). Evaluating accuracy and precision in morphologic traits for sexual dimorphism in malnutrition human skull: a comparative study. International Journal of Morphology. 26(4):877-881.

Gill, G. W. and Gilbert, B. M. (1990). Race identification from the midfacial skeleton: American Blacks and Whites. In Gill, G. W. & S, R. (Eds.). *Skeletal Attribution of Race.* Albuquerque, Maxwell Museum of Anthropology.

Gill, G. W. and Rhine, S. (Eds.) (1990). Skeletal Attribution of Race: Methods for Forensic Anthropology. Anthropological papers of the Maxwell Museum of Anthropology, Number 4.

Gill, G. W. (1995). Challenge on the frontier: discerning American Indians from Whites osteologically. *Journal of Forensic Sciences*, 40: 783-788.

Gill G. W. (1998). Craniofacial criteria in the skeletal attribution of race. In: Reichs K. J., editor. *Forensic osteology: advances in the identification of human remains*. 2nd ed., Springfield: Charles C. Thomas, 293-315.

Gravlee, C. C., Bernard, R. H., Leonard, W. R. (2003). Heredity, environment, and cranial form: a reanalysis of Boas`s immigrant data. *American Anthropologist*, 105(1):125-138.

Gravlee, C. C. (2009). How race becomes biology: embodiment of social inequality. *American Journal of Physical Anthropology*, 139: 47–57.

Gonzalez, P. N., Bernal, V. and Perez, S. I. (2011). Analysis of sexual dimorphism of craniofacial traits using geometric morphometric techniques. *International Journal of Osteoarchaeology*, 21: 82–91.

Gualdi-Russo, E., M. A. Tasca, and P. Brasili. (1999). Scoring of nonmetric cranial traits: A methodological approach. *Journal of Anatomy*, 195:543–550.

Gunz P., Mitteroecker P., and Bookstein F. L. (2005). Semilandmarks in three dimensions. In: Slice, D. E., editor. *Modern Morphometrics in Physical Anthropology*, New York: Kluwer Academic/Plenum Publishers.

Hallgrímsson, B., Zelditch, M. L., Parsons, T. E., Kristensen, E., Young, N. M. and Boyd, S. K. (2007). Morphometrics and Biological Anthropology in the Postgenomic Age. In :Katzenberg, M. A. and Saunders, S. R,. (eds.). *Biological Anthropology of the Human Skeleton, Second Editio*n, John Wiley & Sons, Inc., Hoboken, NJ, USA.

Handley, L. J., Lawson, A. M., Jerome, G., and Francois, B. (2007). Going the distance: human population genetics in a clinal world. *Trends in Genetics*. 23(9):432-439.

Harrison, F. V. (1995). The persistent power of "race" in the cultural and political economy of racism. *Annual Review of Anthropology*, 24:47-74.

Harvati, K. (2009). Into Eurasia: A geometric morphometric re-assessment of the Upper Cave (Zhoukoudian) specimens. *Journal of Human Evolution*, 57: 751-762

Harvati, K. and Weaver, T. D. (2006). Human cranial anatomy and the differential preservation of population history and climate signatures. *The Anatomical Record Part A: Discoveries in Molecular, Cellular, and Evolutionary Biology*, 288A: 1225–1233.

Hennessy, R. J. and Stringer, C. B. (2002). Geometric morphometric study of the regional variation of modern human craniofacial form. *American Journal of Physical Anthropology*, 117: 37–48.

Hefner, J. T. (2009). Cranial Nonmetric Variation and Estimating Ancestry. *Journal of Forensic Sciences*, 54: 985–995.

Howells, W. W. (1973). *Cranial Variation in Man*. Cambridge: Peabody Museum Papers, No. 67.

Howells, W. W. (1996). Howells' craniometric data on the internet. *American Journal of Physical Anthropology*, 101: 441–442.

Hughes, C. E., Juarez, C. A., Hughes, T. L., Galloway, A., Fowler, G. and Chacon, S. (2011). A simulation for exploring the effects of the "trait list" method's subjectivity on consistency and accuracy of ancestry estimations. *Journal of Forensic Sciences*, 56: 1094–1106.

Humphrey, L. T., Dean, M. C. and Stringer, C. B. (1999). Morphological variation in great ape and modern human mandibles. *Journal of Anatomy*, 195: 491–513.

Husmann, P. R. and Samson, D. R. (2011). In the eye of the beholder: sex and race estimation using the human orbital aperture. *Journal of Forensic Sciences*, 56: 1424–1429.

Keita, S. O. Y. (2007) On Meroitic Nubian crania, FORDISC 2.0 and human biological history. *Current Anthropology*, 48:425-427.

Kemkes A. (2007). The unknown female from Cologne: science at a dead end? In: Brickley M. B., Ferllini R., (Eds.). *Forensic anthropology: case studies from Europe*. Springfield: Charles C. Thomas. p. 120–136.

Kendall, D. G. (1977). The diffusion of shape. *Advances in Applied Probability*. 9:428-430.

Kennedy K. A. R. (1995). But Professor, why teach race identification if races don't exist? *Journal of Forensic Science*, 40:797-800.

Kimmerle, E. H., Ross, A. and Slice, D. (2008). Sexual dimorphism in America: geometric morphometric analysis of the craniofacial region. *Journal of Forensic Sciences*, 53: 54–57.

Klepinger, L. L. (2006). *Fundamentals of Forensic Anthropology*. Hoboken: John Wiley and Sons, Inc.

Klingenberg, C. P. and McIntyre, G. S. (1998). Geometric morphometrics of developmental instability: analyzing patterns of fluctuating asymmetry with Procrustes methods. *Evolution*, 52(5):1363-1375.

Klingenberg, C. P., Barluenga, M., and Meyer, A. (2002). Shape analysis of symmetric structures: Quantifying variation among individuals and asymmetry. *Evolution*, 56(10):1909-1920.

Klingenberg, C. P. (2011). MorphoJ: an integrated software package for geometric morphometrics. *Molecular Ecology Resources*, 11:353-357.

Kosiba, S. (2000). Assessing the efficacy and pragmatism of 'race' designation in human skeletal identification: a test of FORDISC 2.0 program. *American Journal of Physical Anthropology*. Supplement 30:200.

Kovarovic, K., Aiello, L. C., Cardini, A., and Lockwood, C. A. (2011). Discriminant function analyses in archaeology: are classification rates too good to be true?. *Journal of Archaeological Science*, 38:3006-3018.

L'Abbe, E. N., Van Rooyen C., Nawrocki S. P., Becker P. J. (2011). An evaluation of non-metric cranial traits used to estimate ancestry in a South African sample. *Forensic Science International*, 209:1-3.

Lele S. (1991) Some comments on coordinate free and scale invariant methods in morphometrics. *American Journal of Physical Anthropology*. 85:407:418.

Lele, S. and Richtsmeier, J. T. (1991a). Euclidean distance matrix analysis: A coordinate-free approach for comparing biological shapes using landmark data. *American Journal of Physical Anthropology*, 86: 415–427.

Lele S. and Richtsmeier J. T. (1991b). On comparing biological shapes: Detection of influential landmarks. *American Journal of Physical Anthropology*. 87:49-65.

Lele, S. (1993). Euclidean Distance Matrix Analysis (EDMA): Estimation of mean form and mean form difference. *Mathematical Geology*, 25(5): 573-602.

Lieberman, L. (2001). How "Caucasoids" got such big crania and why they shrank. From Morton to Rushton. *Current Anthropology*, 42 (1): 69–95.

Livingstone, F. B., and Dobzhansky, T. (1962). On the non-existence of human races. *Current Anthropology*. 3(3):279-281.

Lockeyer, N. (2010). *3D-ID: An assessment of its utility, and an analysis of the potential of 3D geometric morphometrics in ancestry determination from the skull*. B.Sc. Thesis. University of Dundee: U.K.

Loth S. R. and Henneberg M. (1996). Mandibular ramus flexure: a new morphologic indicator of sexual dimorphism in the human skeleton. *American Journal of Physical Anthropology* 99:473–485.

Martin, R. (1928). *Lehrbuch der Anthropologie*. Jena: Gustav Fischer Verlag.

Martin, R. and Saller, K. (1957). *Lehrbuch der Anthropologie*. Stuttgard: Gustav Fischer.

Martínez-Abadías, N., González-José, R., González-Martín, A., Van der Molen, S., Talavera, A., Hernández, P. and Hernández, M. (2006). Phenotypic evolution of human craniofacial morphology after admixture: A geometric morphometrics approach. *American Journal of Physical Anthropology*, 129: 387–398.

Mitteroecker, P., and Gunz, P. (2009). Advances in geometric morphometrics. *Evolutionary Biology* 36(2):235-247.

Molnar, S. (1998). *Human variation: Races, types and ethnic groups* (4th ed.). Upper Saddle River, NJ: Prentice Hall.

Moore-Jansen, P. H., Ousley, S. D., and Jants, R. L. (1994) *Data Collection Procedures for Forensic Skeletal Material*. Forensic Anthropology Center, Department of Anthropology, The University of Tennessee, Knoxville. Report of Investigations No. 48.

Nicholson, E. and Harvati, K. (2006). Quantitative analysis of human mandibular shape using three-dimensional geometric morphometrics. *American Journal of Physical Anthropology*, 131: 368–383

Oettlé, A., Pretorius, E. and Steyn, M. (2005). Geometric morphometric analysis of mandibular ramus flexure. *American Journal of Physical Anthropology*, 128: 623–629.

O'Higgins, P. and Jones, N. (1998). Facial growth in Cercocebus torquatus: an application of three-dimensional geometric morphometric techniques to the study of morphological variation. *Journal of Anatomy*, 193: 251–272.

O'Higgins, P. and Strand- Viðarsdóttir, U. (1999). New approaches to the quantitative analysis of craniofacial growth and variation. In Human growth in the past: studies from bones and teeth. Cambridge: Cambridge

O'Higgins, P. (2000). The study of morphological variation in the hominid fossil record: biology, landmarks and geometry. *Journal of Anatomy*, 197: 103–120.

Ousley, S., Jantz, R. and Freid, D. (2009). Understanding race and human variation: Why forensic anthropologists are good at identifying race. *American Journal of Physical Anthropology*, 139: 68–76.

Ousley, S. D., and Jantz, R. L. (2005). FORDISC 3.0: Personal Computer Forensic Discriminant Functions. University of Tennessee.

Park, H. K., Chung, J. W., and Kho, H. S. (2006). Use of hand-held laser scanning in the assessment of craniometry. *Forensic Science International*, 160(2-3)200-206.

Perez, I. S., Bernal, V., and Gonzalez, P. N. (2006). Differences between sliding semi-landmarks methods in geometric morphometrics, with an applications to human craniofacial and dental variation. *Journal of Anatomy*, 208:769-784.

Pietrusewsky M. (2000). Metric analysis of skeletal remains: methods and applications. In: Katzenburg M. A. and Saunders S. R. (eds.), Biological anthropology of the human skeleton. New York: Wiley-Liss.

Polly, P. D. and Head, J. J. (2004). Maximum-likelihood identification of fossils: taxonomic identification of Quaternary marmots (Rodentia, Mammalia) and identification of vertebral position in the pipe snake *Cylindrophis* (Serpentes, Reptilia). Pps. 197-222, In: A. M. T. Elewa (ed.), *Morphometrics- Applications in Biology and Paleontology*. Springer-Verlag, Berlin, Heidelberg, New York.

Pretorius, E., Steyn, M. and Scholtz, Y. (2006). Investigation into the usability of geometric morphometric analysis in assessment of sexual dimorphism. *American Journal of Physical Anthropology*, 129: 64–70.

Ramsthaler, F., Kettner, M., Gehl, A., and Verhoff M. A. (2010). Digital forensic osteology: morphological sexing of skeletal remains using volume-rendered cranial CT scans. *Forensic Science International*, 195(1-3): 148-152.

Randolph-Quinney, P., Mallett, X. and Black, S. (2009). Anthropology. Wiley Encyclopedia of Forensic Science.

Rhine, S. (1990). Non-metric skull racing. In: Gill, G. W. & Rhine, S. (Eds.) Skeletal Attribution of Race. Maxwell Museum of Anthropology, Anthropological Papers No. 4. p. 9-20.

Richtsmeier, J. T., Paik, C. H., Elfert, P. C., Cole III, T. M., Dahlman, H. R. (1995). Precision, repeatability, and validation of the localization of cranial landmarks using computed tomography scans. *The Cleft Palate-Craniofacial Journal*, 32:217-227.

Richtsmeier, J. T., Burke Deleon, V. and Lele, S. R. (2002). The promise of geometric morphometrics. *American Journal of Physical Anthropology*, 119: 63–91.

Robinson, C., Eisma, R., Morgan, B., Jeffery, A., Graham, E. A. M., Black, S. and Rutty, G. N. (2008). Anthropological Measurement of Lower Limb and Foot Bones Using Multi-Detector Computed Tomography. *Journal of Forensic Sciences*, 53: 1289–1295

Rodríguez-Mendoza, R., Muñoz, M., and Saborido-Rey, F. (2011). Ontogenetic allometry of the bluemouth, Helicolenus dactylopterus dactylopterus (TeleosteiL Scorpaenidae), in the Northeast Atlantic and Mediterranean based on geometric morphometrics. *Hydrobiologia*, 670(1):5-22.

Rogers, T. L. and Allard T. T. (2004). Expert testimony and positive identification of human remains through cranial suture patterns. *Journal of Forensic Sciences*, 49(2): 203-207.

Rohlf, F. J. and Slice, D E. (1990). Extensions of the Procrustes method for the optimal superimposition of landmarks. *Systematic Zoology*, 39:40-59.

Rohlf, F. J. (2003). Bias and error in estimates of mean shape in geometric morphometrics. *Journal of Human Evolution*, 44(6):665-683.

Ross, A., Mckeown, A., and Konisberg, L. (1999). Technical Note: Allocation of crania to groups via the "new morphometry". *Journal of Forensic Sciences*, 43, 584-587.

Ross, A. H., Slice, D. E., and Williams, S. E. (2010). Geometric morphometric tools for the classification of human skulls. *National Criminal Justice Reference Service*, Office of Justice Programs, U.S. Department of Justice, https://www.ncjrs.gov/App/Publications/abstract.aspx?ID=253252. Last accessed 12/12/2011.

Sauer, N. J. (1992). Forensic anthropology and the concept of race: If races don't exist, why are forensic anthropologists so good at identifying them?. Social Science and Medicine, 34(2): 107-111.

Seetah, T. K., Cardini, A., and Miracle, P. T. (2012). Can morphospace shed light on cave bear spatial-temporal variation? Population dynamics of Ursus spelaeus from Romualdova pecina and Vindija, (Croatia). *Journal of Archaeological Science*, 39:500-510.

Shlens, J. (2005). A tutorial on principal component analysis. December 10, 2005, Version 2.

Sholts, S. B., Flores, L., Walker, P. L., and Warmlander, S. K. T. S. (2011a). Comparison of coordinate measurement precision of different landmark types on human crania using a 3D laser scanner and a 3D digitiser: implications for applications of digital morphometrics. *International Journal of Osteoarchaeology*, 21:535-543.

Sholts, S. B., Walker, P. L., Kuzminsky, S. C., Miller, K. W. and Wärmländer, S. K. (2011b). Identification of group affinity from cross-sectional contours of the human midfacial skeleton using digital morphometrics and 3D laser scanning technology. *Journal of Forensic Sciences*, 56: 333–338.

Simmons, T. and W. Haglund. (2005). Anthropology in a Forensic Context. In: Forensic Archaeology: Advances in Theory and Practice. Eds. J. Hunter and M. Cox, Chapter 6, pp. 159-176. Routledge: London.

Singh, N., Harvati, K., Jean-Jacques, H., and Klingenberg, C. P. (2012). Morphological evolution through integration: A quantitative study of cranial integration in Homo, Pan, Gorilla and Pongo. *Journal of Human Evolution*, 62:155-164.

Slice, D. E. (2005). Modern Morphometrics. In D. E. Slice (Ed.), Modern Morphometrics in Physical Anthropology (pp. 1-45). New York: Kluwer Academic/Plenum Publishers.

Slice, D. E. (2007). Geometric Morphometrics. *Annual Review of Anthropology*. 36:261-81.

Slice, D. E. and A. Ross. (2009). 3D-ID: geometric morphometric classification of crania for forensic scientists. http://www.3d-id.org

Spradley, M. K. and Jantz, R. L. (2011). Sex estimation in forensic anthropology: skull versus postcranial elements. *Journal of Forensic Sciences*, 56: 289–296.

Ubelaker, D. H. (1996). Skeletons testify: Anthropology in forensic science—AAPA luncheon address: April 12, 1996. *American Journal of Physical Anthropology*, 101: 229–244.

Ukoha, U., Egqu, O. A., Okafor, I. J., Ogugua, P. C., Onwudinjo, O., and Udemezue, O. O. (2011). Orbital dimensions of adult male Nigerians: a direct measurement study using dry skulls. *International Journal of Biological and Medical Research*, 2(3):688-690.

Viðarsdóttir U. S., O'Higgins P., and Stringer C. A. (2001). Geometric morphometric study of regional differences in the ontogeny of the modern human facial skeleton. *Journal of Anatomy*, 201:211–229

Viscosi V., Cardini, A. (2011). Leaf morphology, taxonomy and geometric morphometrics: a simplified protocol for beginners. *PLoS ONE*, 6(10):1-20.

von Cramon-Taubadel, N., Frazier, B. C. and Lahr, M. M. (2007). The problem of assessing landmark error in geometric morphometrics: Theory, methods, and modifications. *American Journal of Physical Anthropology*, 134:24–35.

Waitzman, A. A., Posnick, J. C., Armstrong, D. C., and Pron, G. E. (1992). Craniofacial skeletal measurements based on computed tomography: Part II. Normal values and growth trends. *The Cleft Palate-Craniofacial Journal*. 29(2):118-128.

Weber, G. W. and Bookstein, F. L. (2011). Virtual Anthropology: A guide to a new interdisciplinary field: Austria: SpringerWienNewYork.

Weber, G. W., Schaefer, K., Prossinger, H., Gunz, P., Mitteroecker, P., and Seidler, H. (2001). Virtual Anthropology: The digital evolution in anthropological sciences. *Journal of Physiological Anthropology and Applied Human Science*, 20(2):69-80.

Wescott D. J., Jantz R. L. (2005). Assessing craniofacial secular change in American Blacks and Whites using geometric morphometry. In Modern Morphometric in Physical Anthropology. Slice, Dennis E. (Ed.), pp. 231–45. Springer U.S.

White, T. D. and Folkens, P. A. (1991). *Human Osteology*. San Diego: Elsevier Academic Press

White, T. D. and Folkens P. A. (2005). *The Human Bone Manual*. Amsterdam: Elsevier Academic Press.

Williams, F. L., and Richtsmeier, J. T. (2003). Comparison of mandibular landmarks from computed tomography and 3D digitizer data. *Clinical Anatomy*, 16:494-500.

Williams, F. L., Belcher, R. L. & Armelagos, G. J. (2005). Forensic misclassification of ancient Nubian crania: implications for assumptions about human variation. *Current Anthropology*, 46: 340–346.

Wilkinson, C. M. (2004). *Forensic Facial Reconstruction*. Cambridge: Cambridge University Press.

Yaroch, L. A. (1996). Shape analysis using the thin-plate spline: Neanderthal cranial shape as an example. *American Journal of Physical Anthropology*, 101:43–89.

Zelditch, M. L., Swiderski, D. L., Sheets, H. D., and Fink W. L. (2004). *Geometric Morphometrics for Biologists: A Primer*. Elsevier, San Diego.

APPENDIX: TABLES

Discriminant analysis (DA) results with associated ancestry biases Tables 1-5

Table 1 Standard Landmarks DAs					
Shape (1st 29PCs=>95% variance)					
Classification Results					
Ancestry group			Predicted Group Membership		Total
			SSA	EUR	
Cross-validated[b]	Count	SSA	27	4	31
		EUR	4	27	31
	%	SSA	87.1	12.9	100.0
		EUR	12.9	87.1	100.0
87.1% of cross-validated grouped cases correctly classified with no associated ancestry bias.					
Size (CS)					
Classification Results					
Ancestry group			Predicted Group Membership		Total
			SSA	EUR	
Cross-validated[b]	Count	SSA	15	16	31
		EUR	11	20	31
	%	SSA	48.4	51.6	100.0
		EUR	35.5	64.5	100.0
56.5% of cross-validated grouped cases correctly classified with an associated ancestry bias of 16.1%.					
Form (PCS of form 1-30)					
Classification Results					
Ancestry group			Predicted Group Membership		Total
			SSA	EUR	
Cross-validated[b]	Count	SSA	26	5	31
		EUR	4	27	31
	%	SSA	83.9	16.1	100.0
		EUR	12.9	87.1	100.0
85.5% of cross-validated grouped cases correctly classified with an associated ancestry bias of 3.2%.					
"size-corrected" (PCS of form 2-30)					
Classification Results					
Ancestry group			Predicted Group Membership		Total
			SSA	EUR	
Cross-validated[b]	Count	SSA	25	6	31
		EUR	2	29	31
	%	SSA	80.6	19.4	100.0
		EUR	6.5	93.5	100.0
87.1% of cross-validated grouped cases correctly classified with an associated ancestry bias of 12.9%.					

Table 2 Vault DAs

Shape (1st 9 PCs=>95% variance)

Classification Results

Ancestry group			Predicted Group Membership		Total
			SSA	EUR	
Cross-validated[b]	Count	SSA	25	6	31
		EUR	9	22	31
	%	SSA	80.6	19.4	100.0
		EUR	29.0	71.0	100.0

75.8% of cross-validated grouped cases correctly classified with an associated ancestry bias of -9.6%.

Size (CS)

Classification Results

Ancestry group			Predicted Group Membership		Total
			SSA	EUR	
Cross-validated[b]	Count	SSA	18	13	31
		EUR	13	18	31
	%	SSA	58.1	41.9	100.0
		EUR	41.9	58.1	100.0

58.1% of cross-validated grouped cases correctly classified with no associated ancestry bias.

Form (PCs of form 1-10)

Classification Results

Ancestry group			Predicted Group Membership		Total
			SSA	EUR	
Cross-validated[b]	Count	SSA	25	6	31
		EUR	9	22	31
	%	SSA	80.6	19.4	100.0
		EUR	29.0	71.0	100.0

75.8% of cross-validated grouped cases correctly classified with an associated ancestry bias of-9.6%.

"size-corrected" (PCs of form 2-10)

Classification Results

Ancestry group			Predicted Group Membership		Total
			SSA	EUR	
Cross-validated[b]	Count	SSA	25	6	31
		EUR	10	21	31
	%	SSA	80.6	19.4	100.0
		EUR	32.3	67.7	100.0

74.2% of cross-validated grouped cases correctly classified with an associated ancestry bias of -12.9%.

Table 3 Orbits Das

Shape (1st 8Pcs=>95% variance)

Classification Results

Ancestry group			Predicted Group Membership		Total
			SSA	EUR	
Cross-validated[b]	Count	SSA	21	10	31
		EUR	8	23	31
	%	SSA	67.7	32.3	100.0
		EUR	25.8	74.2	100.0

71.0% of cross-validated grouped cases correctly classified with an associated ancestry bias of 6.5%.

Size (CS)

Classification Results

Ancestry group			Predicted Group Membership		Total
			SSA	EUR	
Cross-validated[b]	Count	SSA	15	16	31
		EUR	14	17	31
	%	SSA	48.4	51.6	100.0
		EUR	45.2	54.8	100.0

51.6% of cross-validated grouped cases correctly classified with an associated ancestry bias of 6.4%.

Form (PCs of form 1-9)

Classification Results

Ancestry group			Predicted Group Membership		Total
			SSA	EUR	
Cross-validated[b]	Count	SSA	22	9	31
		EUR	7	24	31
	%	SSA	71.0	29.0	100.0
		EUR	22.6	77.5	100.0

74.3% of cross-validated grouped cases correctly classified with an associated ancestry bias of 6.5%.

"size-corrected" (PCs of form 2-9)

Classification Results

Ancestry group			Predicted Group Membership		Total
			SSA	EUR	
Cross-validated[b]	Count	SSA	22	9	31
		EUR	5	26	31
	%	SSA	71.0	29.0	100.0
		EUR	16.1	83.9	100.0

77.5% of cross-validated grouped cases correctly classified with an associated ancestry bias of 12.9%.

Table 4 Nasals Das

Shape (1st 4PCSs =>95%)

Classification Results

Ancestry group			Predicted Group Membership		Total
			SSA	EUR	
Cross-validated[b]	Count	SSA	23	8	31
		EUR	7	24	31
	%	SSA	74.2	25.8	100.0
		EUR	22.6	77.4	100.0

75.8% of cross-validated grouped cases correctly classified with an associated ancestry bias of 3.2%.

Size (CS)

Classification Results

Ancestry group			Predicted Group Membership		Total
			SSA	EUR	
Cross-validated[b]	Count	SSA	23	8	31
		EUR	10	21	31
	%	SSA	74.2	25.8	100.0
		EUR	32.3	67.7	100.0

71.0% of cross-validated grouped cases correctly classified with an associated ancestry bias of -6.5%.

Form (PCs of form 1-5)

Classification Results

Ancestry group			Predicted Group Membership		Total
			SSA	EUR	
Cross-validated[b]	Count	SSA	25	6	31
		EUR	5	26	31
	%	SSA	80.6	19.4	100.0
		EUR	16.1	83.9	100.0

82.3% of cross-validated grouped cases correctly classified with an associated ancestry bias of 3.3%.

"Size-corrected" (PCs of form 2-5)

Classification Results

Ancestry group			Predicted Group Membership		Total
			SSA	EUR	
Cross-validated[b]	Count	SSA	17	14	31
		EUR	15	16	31
	%	SSA	54.8	45.2	100.0
		EUR	48.4	51.6	100.0

53.2% of cross-validated grouped cases correctly classified with an associated ancestry bias of -3.2%.

Table 5 Entire dataset DAs (landmarks and semilandmarks combined)

Shape (1st 29PCs=>95% variance)

Classification Results

Ancestry group			Predicted Group Membership		Total
			SSA	EUR	
Cross-validated[b]	Count	SSA	28	3	31
		EUR	3	28	31
	%	SSA	90.3	9.7	100.0
		EUR	9.7	90.3	100.0

90.3% of cross-validated grouped cases correctly classified with no associated ancestry bias.

Size (CS)

Classification Results[a,c]

Ancestry group			Predicted Group Membership		Total
			SSA	EUR	
Cross-validated[b]	Count	SSA	12	19	31
		EUR	14	17	31
	%	SSA	38.7	61.3	100.0
		EUR	45.2	54.8	100.0

46.8% of cross-validated grouped cases correctly classified with an associated ancestry bias of 16.1%.

Form (PCs of form 1-30)

Classification Results

Ancestry group			Predicted Group Membership		Total
			SSA	EUR	
Cross-validated[b]	Count	SSA	27	4	31
		EUR	3	28	31
	%	SSA	87.1	12.9	100.0
		EUR	9.7	90.3	100.0

88.7% of cross-validated grouped cases correctly classified with an associated ancestry bias of 3.2%.

"**size corrected**" (PCs of form 2-30)

Classification Results

Ancestry group			Predicted Group Membership		Total
			SSA	EUR	
Cross-validated[b]	Count	SSA	28	3	31
		EUR	2	29	31
	%	SSA	90.3	9.7	100.0
		EUR	6.5	93.5	100.0

91.9% of cross-validated grouped cases correctly classified with an associated ancestry bias of 3.2%.

Table 6 Upper, lower, and mean posterior probabilities (PP) for correct assignment from each dataset using all variables

Dataset	Variable	lowerPP(2.5%)	upperPP(97.5%)	meanPP
STANDARD LANDMARKS	size	0.502	0.541	0.520
	shape	0.884	1.000	0.989
	form	0.884	1.000	0.989
	'size-corrected'	0.700	1.000	0.989
VAULT	size	0.509	0.644	0.570
	shape	0.539	0.999	0.842
	form	0.561	0.999	0.838
	'size-corrected	0.532	0.998	0.842
ORBITS	size	0.508	0.701	0.579
	shape	0.558	0.997	0.838
	form	0.562	1.000	0.863
	'size-corrected	0.518	0.998	0.838
NASALS	size	0.531	0.965	0.746
	shape	0.517	0.986	0.841
	form	0.552	0.999	0.889
	'size-corrected	0.502	0.785	0.841
TOTAL CONFIGURATION	size	0.501	0.505	0.504
	shape	0.964	1.000	0.995
	form	0.961	1.000	0.997
	'size-corrected	0.932	1.000	0.995

Table 7 ‘Chance-corrected’ (TAU) and discriminant analysis (DA) accuracy rates for all datasets using all variables

Dataset	Variable	TAU statistic (%)	DA Accuracy rate (%)
Standard landmarks	size	12.9	56.5
	shape	74.2	87.1
	form	71.0	87.1
	‘size-corrected’	74.2	87.1
Vault	size	16.1	58.1
	shape	51.6	75.8
	form	51.6	75.8
	‘size-corrected’	48.4	74.2
Orbits	size	3.2	51.6
	shape	41.9	71.0
	form	48.4	74.2
	‘size-corrected’	54.8	77.4
Nasals	size	41.9	71.0
	shape	51.6	75.8
	form	64.5	82.3
	‘size-corrected’	6.5	53.2
Total configuration	size	-6.5	46.8
	shape	80.6	90.3
	form	77.4	88.7
	‘size-corrected’	83.9	91.9